CANADIAN JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. XI - NO. 4
1959

THE UNIVERSITY OF MICHIGAN
1959
MATHEMATICS

Complemented modular lattices	Ichiro Amemiya and Israel Halperin	481
Sur un théorème de G. Szekeres, concernant les fonctions monotones et convexes	S. Marcus	521
Some results on uniqueness and successive approximations	Fred Brauer	527
The asymptotic behaviour of the Laurent coefficients	Max Wyman	534
Convex structures and continuous selections	Ernest Michael	556
A note on hyperconvexity in Riemannian manifolds	Albert Nijenhuis	576
The a-local operator problem	Louis de Branges	583
On a theorem of Hayman concerning quasi-bounded functions	P. B. Kennedy	593
Some principles underlying the construction of measures	Charles A. Hayes, Jr.	601
Pencils of null polarities	Seymour Schuster	614
On the symmetry of cubic graphs	W. T. Tutte	621
Quadrics over GF(2) and their relevance for the cubic surface group	W. L. Edge	625
Polytopes over GF(2) and their relevance for the cubic surface group	H. S. M. Coxeter	646
On the structure of half-groups	James V. Whittaker	651
On the class number of representations of an order	Irving Reiner	660

Published for
THE CANADIAN MATHEMATICAL CONGRESS
by the
University of Toronto Press

EDITORIAL BOARD

H. S. M. Coxeter, G. F. D. Duff, R. D. James, R. L. Jeffery, J.-M. Maranda, G. de B. Robinson, H. Zassenhaus

with the co-operation of

A. D. Alexandrov, R. Brauer, W. P. Brown, D. B. DeLury, J. Dixmier, P. Hall, N. S. Mendelsohn, P. Scherk, J. L. Synge, A. W. Tucker, W. J. Webber, M. Wyman

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, G. F. D. Duff, University of Toronto. Authors are asked to write with a sense of perspective and as clearly as possible, especially in the introduction. Regarding typographical conventions, attention is drawn to the Author's Manual of which a copy will be furnished on request.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is $8.00. This is reduced to $4.00 for individual members of recognized Mathematical Societies.

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of Alberta
University of British Columbia
Dalhousie University
Université Laval
University of Manitoba
McMaster University
Queen's University
St. Mary's University
Assumption University
Carleton College
Ecole Polytechnique
Loyola College
McGill University
Université de Montréal
Royal Military College
University of Toronto

National Research Council of Canada
and the
American Mathematical Society

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

COMPLEMENTED MODULAR LATTICES

ICHIRO AMEMIYA[1] AND ISRAEL HALPERIN

1. Introduction.

1.1. This paper gives a lattice theoretic investigation of "finiteness"[2] and "continuity of the lattice operations" in a complemented modular lattice. Although we usually assume that the lattice is $\aleph$-complete for some infinite $\aleph$,[3] we do not require completeness and continuity, as von Neumann does in his classical memoir on continuous geometry **(3)**; nor do we assume orthocomplementation as Kaplansky does in his remarkable paper **(1)**.

1.2. Our exposition is elementary in the sense that it can be read without reference to the literature. Our brief preliminary § 2 should enable the reader to read this paper independently.

1.3. Von Neumann's theory of independence **(3**, Part I, Chapter II**)** leans heavily on the assumption that the lattice is continuous, or at least upper continuous. We do not assume such continuity and we find it necessary therefore to distinguish several concepts of independence for a family of elements $\{a_\lambda; \lambda \in \Lambda\}$: *independence* shall mean that $a_\lambda \sum(a_\mu; \mu \in F) = 0$ whenever F is a *finite* subset of Λ and $\lambda \notin F$; *residual independence* shall mean that $a_\lambda \sum(a_\mu; \mu \neq \lambda) = 0$ for every λ; and *strong independence* shall mean that $\prod_\lambda \sum(a_\mu; \mu \neq \lambda) = 0$.

Strong independence is sufficiently restrictive that, even without assuming continuity of the lattice operations, many of the continuous geometry arguments remain vaild. For example, if $\{a_\lambda, b_\lambda; \lambda \in \Lambda\}$ is strongly independent and for each λ there is given a perspective mapping of $[0, a_\lambda]$ onto $[0, b_\lambda]$, then these mappings can be imbedded in a single perspective mapping of $[0, \sum a_\lambda]$ onto $[0, \sum b_\lambda]$.

§ 3 is devoted to a discussion of independence.

1.4. Suppose L is complemented, modular, and *countably complete*. Von Neumann's arguments **(3**, Part I, Theorem 4.3**)** show that L is finite, that is,

Received April 29, 1958; presented to the American Mathematical Society, January 29, 1958.

[1]Post-doctoral Fellow (of the National Research Council of Canada) at Queen's University, on leave of absence from Tokyo College of Science.

[2]In this paper, a lattice is called finite if every independent sequence of pairwise perspective elements has all its elements coinciding with zero. Sequence shall mean *infinite* sequence throughout this paper. We note however that finite families $\{a_1, \ldots, a_m\}$ of pairwise perspective elements were used by von Neumann in his coordinatization theory **(3**, Part II, chapter III**)** and play a key role there.

[3]Throughout this paper $\aleph$ denotes an arbitrary (but fixed) infinite cardinal (that is $\aleph > \aleph_0$); Ω denotes the least ordinal number whose corresponding cardinal power is $\aleph$.

an independent sequence $\{a_n\}$ of pairwise perspective elements with $a_1 \neq 0$ *cannot* exist, if the lattice is $\aleph_0$-continuous (this means: the lattice is both upper $\aleph_0$-continuous and lower $\aleph_0$-continuous).

If $\aleph_0$-continuity does not hold, then such sequences $\{a_n\}$ can occur. But we find the paradoxical result: *the existence of such sequences actually forces a certain type of continuity to hold.* This situation is described more precisely in the following paragraph.

A *homogeneous* sequence is defined to be a strongly independent sequence $\{a_n\}$ of pairwise perspective elements. We draw attention to two important special cases:

(i) Type (A): all a_n have a common complement, that is, for some element A, $a_n \oplus A = 1$.

(ii) Type (A*): all $a^*_n = \sum(a_m; m \neq n)$ have a common complement.[4]

In § 5 we show: Suppose $\{a_n\}$ is a homogeneous sequence; then the lattice $[0, \sum a_n]$ is upper $\aleph_0$-continuous if and only if $\{a_n\}$ is of type (A), lower $\aleph_0$-continuous if and only if $\{a_n\}$ is of type (A*). Thus, if $\{a_n\}$ is of *both* types (A) and (A*), the above-mentioned result of von Neumann shows that all a_n must be 0.

In § 8 we show that the types (A) and (A*) are *mutually exclusive* in a stronger sense, namely: If $\{a_n\}$ and $\{b_n\}$ are homogeneous sequences of types (A) and (A*) respectively, then $\sum a_n$ and $\sum b_n$ are completely disjoint (this means: a perspective to b with $a \leqslant \sum a_n$ and $b \leqslant \sum b_n$ can occur only when $a = b = 0$). On the other hand, these two types are *exhaustive* in the following sense: every homogeneous sequence $\{a_n\}$ has a unique decomposition $a_n = b_n + c_n$ with $\{b_n\}$ a homogeneous sequence of type (A) and (c_n) a homogeneous sequence of type (A*).

From these facts about homogeneous sequences we can deduce (see § 8): If L is complete then L has a direct sum decomposition $L = L_1 + L_2 + L_3$ where $L_i = (0, a_i)$ with each a_i in the centre of L, and with L_1 upper $\aleph_0$-continuous, L_2 lower $\aleph_0$-continuous and L_3 finite.

1.5. Now suppose L is even $\aleph$-complete for a given infinite $\aleph$. We call L *locally $\aleph$-continuous* if for every $a \neq 0$ there exists some $0 \neq a_1 \leqslant a$ with $[0, a_1]$ $\aleph$-continuous. We show (see Corollary 1 to Theorem 7.1): If L is locally $\aleph$-continuous and finite then L must be $\aleph$-continuous.

1.6. In §§ 4, 6 we establish, among other properties of finiteness and continuity, that they are *additive*, that is, if $[0, a]$ and $[0, b]$ enjoys one of these properties then so does $[0, a + b]$.

1.7. Finally, in § 9 we prove theorems somewhat more general than that of Kaplansky **(1)**. Kaplansky proved: (i) every countably complete ortho-

[4] If $(a_1, \ldots, a_m)$ is a finite independent family of pairwise perspective elements, then (i) and (ii) both hold, that is, the a_n have a common complement and the a^*_n have a common complement.

complemented modular lattice is finite and (ii) every complete orthocomplemented modular lattice is necessarily continuous.

Our work gives lattice theoretic proofs for generalizations of both of these results. In particular, (ii) is strengthened to (ii)′ every $\aleph$-complete orthocomplemented modular lattice is $\aleph$-continuous.

More generally, we prove, generalizing (i):

THEOREM 9.1. *A countably complete, complemented modular lattice is finite, if* (*): *for every* $a \neq 0$ *there exists an anti-automorphism* ϕ *of* L *such that* b *perspective to a subelement of* a *occurs for some* $b \neq 0$ *with* $b\phi(a) = 0$.

(*) holds, for example, if L possesses an orthocomplementation, or even, in the case that L is complete, if L possesses an anti-automorphism which is an orthocomplementation on the centre of L (see Corollary 1 to Theorem 9.1).

We prove, generalizing (ii)′:

THEOREM 9.5. *An* $\aleph$*-complete complemented modular lattice is* $\aleph$*-continuous if it is finite and possesses an anti-automorphism* ϕ *of period two with the following continuity property:* (**) *for every limit ordinal number* $\Omega_1 \leqslant \Omega$, $x_\beta + \phi(x_\beta) = 1$, $x_\beta\phi(x_\beta) = 0$ *for all* $\beta < \Omega_1$ *and* $x_\beta \leqslant x_\gamma$ *for all* $\beta \leqslant \gamma < \Omega_1$ *together imply* $(\sum x_\beta) + \phi(\sum x_\beta) = 1$, $(\sum x_\beta)\phi(\sum x_\beta) = 0$.

Clearly every orthocomplementation ϕ has the property (**).

1.8. An alternative (but still lattice theoretic) proof of the Kaplansky's finiteness theorem for the *orthocomplemented* case (see (i) in § 1.7 above) is given in an Appendix. This Appendix can be read independently of the rest of this paper and it is somewhat related to Kaplansky's original method.

2. Preliminaries.[5]

2.1. Let L be a set of elements partially ordered by a relation $a \leqslant b$ (written also $b \geqslant a$). By definition, partial ordering means: $a \leqslant b$, $b \leqslant c$ imply $a \leqslant c$, and $a \leqslant b$, $b \leqslant a$ hold if and only if $a = b$ (that is, a and b are the same element).

When a_λ is in L for each $\lambda \in \Lambda$ we call a the *union* of the a_λ and write $a = \sum_{\lambda\in\Lambda} a_\lambda$ (or $\sum a_\lambda$) if a is an element such that: $x \geqslant a_\lambda$ for every λ is equivalent to $x \geqslant a$. We call a the *meet* of the a_λ and write $a = \prod_{\lambda\in\Lambda} a_\lambda$ (or $\prod a_\lambda$) if a is an element such that: $x \leqslant a_\lambda$ for every λ is equivalent to $x \leqslant a$ (each of union and meet is clearly unique if it exists at all).

The zero (unit) in L written as $0(1)$, is defined to be the element (if it exists) such that $0 \leqslant x (x \leqslant 1)$ holds for all x in L.

The dual to any statement or construction concerning elements of L is obtained by replacing $\leqslant$ by $\geqslant$; $\sum, \prod$ by $\prod, \sum$ respectively and 0, 1 by 1, 0, respectively. L' denotes the partially ordered set dual to L. Any theorem implies its dual.

[5] This section is mostly based on the original material of J. von Neumann (see (3) or (2)).

L is called complete if $\sum a_\lambda, \prod a_\lambda$ exist for all families $\{a_\lambda; \lambda \in \Lambda\}$; $\aleph$-*complete* if these elements exist whenever $\bar{\Lambda} < \aleph$;[6] a *lattice* if it is 2-complete (hence n-complete for every $n = 2, 3, \ldots$).

A lattice is called modular if $a(b + c) = b + ac$ whenever $a \geqslant b$ (equivalently, if: $a(b + c) = a(b(a + c) + c)$ for all a, b, c).

When $a \leqslant b$ we write $L(a, b)$ or $[a, b]$ to denote the sub-partially-ordered set of all x with $a \leqslant x \leqslant b$; clearly, it has a, b as zero and unit respectively.

2.2. *Let L be a lattice with zero element.* Elements a, b are called *disjoint* if $ab = 0$ ($\oplus$ shall mean $+$ but shall imply that the summands are disjoint).

If $a \leqslant c$, $[c - a]$ will denote any element A, to be called a *complement of a in c* (sometimes called a *relative complement* of a in c), for which $a \oplus A = c$. If L has a unit, $[1 - a]$ (if it exists) is called a *complement* of a.

L is called *complemented* if 0, 1 exist in L and every a has at least one complement. L is called *orthocomplemented* if 0, 1 exist in L and L possesses an anti-automorphism ϕ of period 2 with $\phi(a) \oplus a = 1$ for all a.

If L is complemented and modular, a relative complement $[c - a]$ exists always ($c[1 - a]$ will do); then, whenever $ab = 0$ there exists a complement A of a with $A \geqslant b$ (indeed, $b + [1 - (a + b)]$ will do for A).

If L is modular and A is a complement of a then $[0, a]$ and $[A, 1]$ are lattice isomorphic under the mutually inverse mappings:

$$a_1 \to a_1 + A \text{ if } a_1 \leqslant a; \quad A_1 \to aA_1 \text{ if } A_1 \geqslant A. \tag{2.1}$$

2.3. *Let L be a modular lattice with zero element.* The elements a and b are called *perspective* with axis x (we write $a \sim b$), if $a \oplus x = b \oplus x$; we may replace x by $x(a + b)$ to obtain $a \oplus x = b \oplus x = a + b$.

If a, b are perspective with axis x then $[0, a]$ and $[0, b]$ are lattice isomorphic under the mutually inverse perspective mappings:

$$a_1 \to (a_1 + x)b \text{ if } a_1 \leqslant a; \quad b_1 \to (b_1 + x)a \text{ if } b_1 \leqslant b$$

(clearly, $a_1 \sim b_1$ with the same axis x). We note:

$$a \sim c,\ c \sim b,\ (a + c)b = 0 \text{ imply } a \sim b; \tag{2.2}$$

for if $a \oplus x = x \oplus c = a + c$ and $c \oplus y = y \oplus b = c + b$ then

$$a \oplus (x + y)(a + b) = b \oplus (x + y)(a + b).$$

Elements a, b are called *projective* (we write $a \asymp b$) if $a = a_1$ and $b = a_m$ for some finite family $a_1, \ldots, a_m$ with $a_i \sim a_{i+1}$ for $i < m$.

We shall say that an element a can be *doubled in L* if

$$a \sim u \text{ holds for some } u \text{ in } L \text{ with } ua = 0. \tag{2.3}$$

If a modular lattice L with zero has a unit, we shall say the lattice L *can*

[6] $\bar{\Lambda}$ denotes the cardinal power of Λ. For every family of lattice elements $\{a_\lambda; \lambda \in \Lambda\}$ which we consider, we shall suppose that $\bar{\Lambda} < \aleph$. Von Neumann defines L to be an $\aleph$-lattice if it is $\aleph_1$-complete for every $\aleph_1 < \aleph$ (see (3, Part III, Definition A. 1)).

be doubled if there exists a modular lattice L_1 with zero such that for some u in L_1, L is lattice isomorphic to $[0, u]$ and u can be doubled in L_1. Clearly if a can be doubled in L, the lattice $[0, a]$ can be doubled.

We write (as in von Neumann (**3**, Part II, Definition 3.4)), $(a, b, x)C$ to mean: $a \oplus x = b \oplus x = a \oplus b$.

We call a and b *completely disjoint* and write $(a, b)P$ to mean:

(2.4) $a_1 \sim b_1$, $a_1 \leqslant a$, $b_1 \leqslant b$ together imply $a_1 = 0$.

Clearly $(a, b)P$ implies $ab = 0$.

We say:

(2.5) a is in the *centre* of L if $(a, b)P$ holds whenever $ab = 0$.

2.4. *Let L be a complemented modular lattice.* Then $(a, b)P$ holds if and only if:

(2.6) every complement of b contains a.

(Suppose (2.6) fails: if B is a complement of b with $B \geqslant a$ false, then $a_1 = [a - aB] \neq 0$ and $a_1 \sim (B + a_1)b$ (with axis B) so $(a, b)P$ does not hold. Suppose, on the other hand, (2.6) does hold: then if $a_1 \leqslant a$, $b_1 \leqslant b$, and $a_1 \sim b_1$ with axis x, we have in succession: $a_1 b \leqslant ab = 0$; $b(a_1 + b_1)x = 0$; there exists a complement B of b with $B \geqslant (a_1 + b_1)x$; $B \geqslant (a_1 + b_1)x + a_1$; $B \geqslant b_1$; $b_1 = 0$; $a_1 = 0$; hence $(a, b)P$ holds.)

(2.6) is also equivalent to: every complement of a contains b (consequently, a is in the centre of L if and only if it has a unique complement, necessarily also in the centre of L, and a is in the centre of L if and only if it is in the centre of L').

Hence in a complemented modular lattice, if $(a, b_\lambda)P$ holds for every λ, and $\sum b_\lambda$ exists, then every complement of a contains $\sum b_\lambda$ along with all b_λ so $(a, \sum b_\lambda)P$ holds; therefore, if $\sum a_\lambda$ and $\sum b_\mu$ both exist and $(\sum a_\lambda, \sum b_\mu)P$ is false, we must have $(a_\lambda, b_\mu)P$ false for some particular λ, μ.

Consequently, although this fact is not needed in the present paper, if b_λ are all in the centre of L then $\sum b_\lambda$, if it exists, is also in the centre of L and, by duality, $\prod b_\lambda$, if it exists, is also in the centre of L.

If, in a complemented modular lattice, $(a, b)P$ is false and $b \sim c$ then $(a, c)P$ is also false; this follows from:

(2.7) $a \sim b$, $b \sim c$, $a \neq 0$ together imply $a_1 \sim c_1$ for some $a_1 \leqslant a$, $c_1 \leqslant c$ with $a_1 \neq 0$.

Clearly, we need prove (2.7) only for the case $ac = ba = bc = 0$. Because of (2.2), we may also suppose $a \leqslant b + c$, $c \leqslant a + b$.[7] Now it follows that $a \oplus b = c \oplus b$ so $a \sim c$ (axis b).

Hence, in a complemented modular lattice, $(a, b)P$ holds if and only if

[7]For example, if $a \leqslant b + c$ is false then $a' = [a - a(b + c)] \neq 0$, and $a' \sim b'$, $b' \sim c'$ for some $b' \leqslant b$ and $c' \leqslant c$. Since $a'(b' + c') \leqslant a'(b + c) = 0$, it follows from (2.2) that $a' \sim c'$.

$ac = 0$ whenever $c \sim b$. Indeed, $(a, b)P$ and $b \sim c$ imply $ac = 0$ by (2.7). On the other hand, if $ac = 0$ for all $c \sim b$ then $(a, b)P$ does hold; for then $ab = 0$, and if $a_1 \leqslant a$, $b_1 \leqslant b$ with $a_1 \sim b_1$ we have $b \sim (a_1 + [b - b_1])$ (use (2.2)); since $a(a_1 + [b - b_1]) = a_1$, then we must have $a_1 = 0$, proving $(a, b)P$ does hold.

If L is complete then for each element a there exists a central element $z \geqslant a$ (namely, $z = \sum a'$ for all a' perspective to subelements of a) such that:

(2.8) $(a, b)P$ holds if and only if $zb = 0$. This z is the least element in the centre with property $z \geqslant a$.

2.5. *Let L be an $\aleph$-complete lattice.* A family $\{a_\lambda; \lambda < \Omega_1\}$ with $\Omega_1 \leqslant \Omega$, either increasing, that is, $\lambda \leqslant \mu$ implies $a_\lambda \leqslant a_\mu$ (written $a_\lambda \uparrow a$ to denote also $a = \sum a_\lambda$) or decreasing, that is, $\lambda \leqslant \mu$ implies $a_\lambda \geqslant a_\mu$ (written $a_\lambda \downarrow a$ to denote also $a = \prod a_\lambda$) is said to *converge continuously* if for every x, $\sum(xa_\lambda) = xa$ or $\prod(x + a_\lambda) = x + a$, respectively; L is said to be *upper $\aleph$-continuous* or *lower $\aleph$-continuous* if every such increasing or decreasing family respectively, converges continuously (upper $\aleph$-continuity of L is clearly equivalent to lower $\aleph$-continuity of L'.[8]

If $a_\lambda \uparrow$ continuously, then for every c, $ca_\lambda \uparrow$ continuously; indeed, for every x, $x(\sum ca_\lambda) = cx(\sum a_\lambda) = \sum(cx)a_\lambda = \sum x(ca_\lambda)$.

L is called $\aleph$-continuous if it is both upper and lower $\aleph$-continuous.

2.6. *Let L be a complemented, modular, and $\aleph$-complete lattice.* If $\{a_\lambda\}$ is increasing or decreasing, then $\{a_\lambda\}$ does converge continuously if: $xa_\lambda = 0$ for every λ implies $x\sum a_\lambda = 0$ or if $x + a_\lambda = 1$ for every λ impiles $x + \prod a_\lambda = 1$, respectively. Also, L is upper $\aleph$-continuous if $a_\lambda \uparrow 1$ implies a_λ converges continuously, lower $\aleph$-continuous if $a_\lambda \downarrow 0$ implies a_λ converges continuously.

2.7. *Let L be an $\aleph$-complete lattice with zero.* L is called *locally $\aleph$-continuous* (*upper $\aleph$-continuous, lower $\aleph$-continuous*) if $a \neq 0$ implies $[0, a_1]$ is $\aleph$-continuous (upper $\aleph$-continuous, lower $\aleph$-continuous) for some $0 \neq a_1 \leqslant a$.

If L is also complemented and modular, then L is locally $\aleph$-continuous (upper $\aleph$-continuous, lower $\aleph$-continuous) if and only if the dual L' is locally $\aleph$-continuous (lower $\aleph$-continuous, upper $\aleph$-continuous); for if $A \neq 1$, let a be a complement of A. Then $a \neq 0$, and $[0, a_1]$ is $\aleph$-continuous (upper $\aleph$-continuous, lower $\aleph$-continuous) for some $0 \neq a_1 \leqslant a$. Let $A_1 = A + [a - a_1]$. Then $A \leqslant A_1 \neq 1$ and $[A_1, 1]$ is $\aleph$-continuous (upper $\aleph$-continuous, lower $\aleph$-continuous) by (2.1). This shows that L' is locally $\aleph$-continuous (lower $\aleph$-continuous, upper $\aleph$-continuous) since $\leqslant$ in L means $\geqslant$ in L'.

---

[8] Transfinite induction shows that L is upper $\aleph$-continuous if and only if, for arbitrary $\{a_\lambda; \lambda \in \Lambda\}$ and for every x, $\Sigma_F(x\Sigma(a_\lambda; \lambda \in F))$ exists and equals $x\Sigma a_\lambda$ (F varies over all finite subsets of Λ), lower $\aleph$-continuous if and only if, for arbitrary $\{a_\lambda; \lambda \in \Lambda\}$ and for every x, $\Pi_F(x + \Pi(a_\lambda; \lambda \in F))$ exists and equals $x + \Pi a_\lambda$.

An equivalent definition of continuity in terms of *directed* families $\{a_\lambda\}$ was given by U. Sasaki who used a lemma of T. Iwamura (see **(4)** or **(2**, Appendix II**)**).

If $[0, a_\lambda]$ is $\aleph$-continuous (upper $\aleph$-continuous, lower $\aleph$-continuous) for every λ then $[0, \sum a_\lambda]$ must be locally $\aleph$-continuous (upper $\aleph$-continuous, lower $\aleph$-continuous); for if $x \neq 0$ and $x \leqslant \sum a_\lambda$ then $(x, a_\lambda)P$ is false for some λ, so $[0, x_1]$ is lattice isomorphic to $[0, a_\lambda']$ for some $a_\lambda' \leqslant a_\lambda$ and some $x_1 \neq 0$ with $x_1 \leqslant x$. But $[0, a_\lambda']$ is $\aleph$-continuous (upper $\aleph$-continuous, lower $\aleph$-continuous) along with $[0, a_\lambda]$, so $[0, x_1]$ has the same property. This proves that $[0, \sum a_\lambda]$ is locally $\aleph$-continuous (upper $\aleph$-continuous, lower $\aleph$-continuous).

3. Independence theory. In this section we assume L is an $\aleph$-complete modular lattice with zero. Since we do not make any continuity assumptions we need to refine the von Neumann independence theory.[9] in so far as it applies to infinite families of elements. In particular, in Theorem 3.1 below, we use a complementation argument to replace the usual "continuity" argument.

If $\{a_\lambda; \lambda \in \Lambda\}$ is a set of elements in L we use the following notation:

a_λ^* denotes $\sum(a_\mu; \mu \neq \lambda)$,

a_Γ denotes $\sum(a_\lambda; \lambda \in \Gamma)$ if $\Gamma \subset \Lambda$,

a_Γ^* denotes $\sum(a_\lambda; \lambda \notin \Gamma)$ if $\Gamma \subset \Lambda$ (in particular, $a_\Lambda^* = 0$).

Definition 3.1. A family $\{a_\lambda; \lambda \in \Lambda\}$ is called *independent* if $a_\lambda a_F = 0$ whenever F is a finite subset of Λ and $\lambda \notin F$; *residually independent* if $a_\lambda a_\lambda^* = 0$ for every λ.

Definition 3.2. If $\{a_\lambda\}$ is residually independent the residual element of $\{a_\lambda\}$ is defined to be $\prod a_\lambda^*$; an element x is called a *residual element in* L (more precisely, *an* $\aleph$*-residual element in* L) if x is the residual element of some residually independent family $\{a_\lambda; \lambda \in \Lambda\}$ with $\bar{\Lambda} = \aleph$.

If $\{a_\lambda\}$ is residually independent with residual element 0 then $\{a_\lambda\}$ is called *strongly independent*.

Because of the modular law, the following statements follow easily:

Independence of $\{a_\lambda\}$ is equivalent to: $a_F a_G = 0$ whenever F, G are finite, disjoint subsets of Λ, and also to:

$$\prod_\gamma \sum(a_\lambda; \lambda \in F_\gamma) = \sum(a_\lambda; \lambda \in \bigcap F_\gamma)$$

for every *finite* collection of finite subsets F_γ of Λ.

Residual independence implies independence and is equivalent to: $a_F a_F^* = 0$ for every finite subset F of Λ.

Strong independence implies residual independence and is equivalent to the single condition $\Pi a^* = 0$. (It will follow from Theorem 3.1 below that strong independence of $\{a_\lambda\}$ is equivalent to: for every collection of subsets I_γ of Λ, $\prod_\gamma \sum(a_\lambda; \lambda \in I_\gamma)$ exists and equals $\sum(a_\lambda; \lambda \in \bigcap I_\gamma)$.)

[9] The von Neumann theory of independence can be found in (3, Part I, chapter II).

If $a_i(a_1 + \ldots + a_{i-1}) = 0$ for $i > 2$, then the finite or infinite family $\{a_n; n > 1\}$ is independent. Hence:

if $\{b_1, \ldots, b_m, a_1, \ldots, a_r\}$ *is independent and* $c_1 + \ldots + c_p < b_1 + \ldots + b_m$ *and* $\{c_1, \ldots, c_p\}$ *is independent, then* $\{c_1, \ldots, c_p, a_1, \ldots, a_r\}$ *is independent*; a generalization of this fact is proved in the Corollary to Theorem 3.1.

If $b_\lambda < a_\lambda$ for each λ and $\{a_\lambda\}$ is independent (residually independent, strongly independent), then $\{b_\lambda\}$ is independent (residually independent, strongly independent).

If $\{b_\lambda, c_\lambda; \lambda \in \Lambda\}$ is independent (residually independent, strongly independent) then $\{b_\lambda + c_\lambda\}$ is independent (residually independent, strongly independent).

$\{a_n; n = 1, 2, \ldots\}$ is residually independent if and only if $a_n \sum (a_m; m > n) = 0$ for every $n = 1, 2, \ldots$, strongly independent if and only if residually independent with $\prod_n(\sum(a_m; m > n)) = 0$. If $\{a_n\}$ is strongly independent, then

$$c_n < \sum_{i=n}^{\infty} a_i, c_n\left(\sum_{i=n+1}^{\infty} a_i\right) = 0 \tag{3.1}$$

for every n imply $\{c_n\}$ is strongly independent.

If $\{a_\lambda\},\{b_\lambda\}$ are both independent (residually independent) and $(\sum a_\lambda)(\sum b_\lambda) = 0$ then $\{a_\lambda + b_\lambda; \lambda \in \Lambda\}$ and $\{a_\lambda, b_\lambda; \lambda \in \Lambda\}$ are both independent (residually independent) (the Corollary to Theorem 3.1 below shows that if $\{a_\lambda\},\{b_\lambda\}$ are both strongly independent with $(\sum a_\lambda)(\sum b_\lambda) = 0$ then $\{a_\lambda + b_\lambda; \lambda \in \Lambda\}$ and $\{a_\lambda, b_\lambda; \lambda \in \Lambda\}$ are both strongly independent).

If L is upper $\aleph$-continuous then independence implies strong independence for families $\{a_\lambda; \lambda \in \Lambda\}$ with $\bar{\Lambda} < \aleph$ (this was shown first by von Neumann (3, Part I, Chapter II)).

THEOREM 3.1. *Suppose* $\{a_\lambda\}$ *is strongly independent and for an arbitrary set of* μ, $a_{\lambda,\mu} < a_\lambda$ *for all* λ, μ. *Then* $\prod_\mu a_{\lambda,\mu}$ *exists for each* λ *if* $\prod_\mu(\sum_\lambda a_{\lambda,\mu})$ *exists. On the other hand,*

$$\prod_\mu(\sum_\lambda a_{\lambda,\mu}) = \sum_\lambda(\prod_\mu a_{\lambda,\mu})$$

(that is, both sides exist and are equal) provided that for each λ, *the element* $\prod_\mu a_{\lambda,\mu}$ *exists and has a complement in* a_λ *(in particular, if* $\prod_\mu a_{\lambda,\mu} = 0$ *for every* λ*).*

Proof. 1. If $\prod_\mu a_{\lambda,\mu}$ exists let it be denoted as b_λ.

Clearly, if $\prod_\mu(\sum_\lambda a_{\lambda,\mu})$ exists, then for each ν,

$$a_\nu \prod_\mu(\sum_\lambda a_{\lambda,\mu}) = \prod_\mu(a_{\nu,\mu} + 0)$$

since $\{a_\lambda\}$ is strongly independent, so b_ν exists for each ν.

2. Clearly $\sum_\lambda a_{\lambda,\mu} > \sum_\lambda b_\lambda$ for every μ. So we need only show: if for some x, $\sum_\lambda a_{\lambda,\mu} > x$ for all μ, then $x < \sum_\lambda b_\lambda$.

But if for all μ, $\sum_\lambda a_{\lambda,\mu} \geqslant x$ then for all λ, μ,

$$x \leqslant a_{\lambda,\mu} + a^*_\lambda,$$
$$(x + a^*_\lambda)a_\lambda \leqslant (a_{\lambda,\mu} + a^*_\lambda)a_\lambda = a_{\lambda,\mu};$$

then for every λ,

$$(x + a^*_\lambda)a_\lambda \leqslant \prod_\mu a_{\lambda,\mu} = b_\lambda,$$
$$x \leqslant x + a^*_\lambda \leqslant b_\lambda + a^*_\lambda;$$

finally,

$$x \leqslant \prod_\lambda (b_\lambda + a^*_\lambda).$$

Thus the theorem will be completely proved if we establish

$$\prod_\lambda (b_\lambda + a^*_\lambda) = \sum_\lambda b_\lambda.$$

3. Now suppose $c_\lambda \oplus b_\lambda = a_\lambda$ for each λ. Then

$$\sum b_\lambda + \sum c_\lambda \geqslant \prod (b_\lambda + a^*_\lambda) \geqslant \sum b_\lambda,$$

and the modular law now shows that we need only prove

$$(\prod (b_\lambda + a^*_\lambda))(\sum c_\lambda) = 0.$$

And this does hold because $(b_\lambda + a^*_\lambda)(\sum c_\lambda) \leqslant a^*_\lambda$ for each λ, and $\prod a^*_\lambda = 0$.

COROLLARY. (*Strong independence under substitution*). *Suppose L is an $\aleph$-complete modular lattice with zero, $\{a_\lambda; \lambda \in \Lambda\}$ is strongly independent and $\Gamma, \Delta, \ldots$ are mutually disjoint subsets of Λ. If $\{c_\mu\}, \{d_\nu\}, \ldots$ are each strongly independent with sets of indices $\mu, \nu, \ldots$ each of cardinal power $< \aleph$ and $a_\Gamma \geqslant \sum c_\mu$, $a_\Delta \geqslant \sum d_\nu, \ldots$, then the set of all elements $\{\text{all } c_\mu, \text{ all } d_\nu, \ldots\}$ is strongly independent.*

Proof. Since $a^*_\Gamma \leqslant \prod (a^*_\lambda; \lambda \in \Gamma)$ the meet $a^*_\Gamma a^*_\Delta \ldots \leqslant \prod (a^*_\lambda; \lambda \in \Lambda) = 0$. So $\{a_\Gamma, a_\Delta, \ldots\}$ is strongly independent.

To prove $\{\text{all } c_\mu, \text{ all } d_\nu, \ldots\}$ strongly independent we form the union $\sum^*$ of *all c_μ, all d_ν, omitting one of these elements* and we need only prove that all such $\sum^*$ have 0 as meet.

But if c_μ is omitted, $\sum^* = c^*_\mu + a^*_\Gamma$. Then Theorem 3.1, applied to the family $\{a_\Gamma, a_\Delta, \ldots\}$ shows that

$$\prod \sum^* = \prod_\mu c^*_\mu + \prod_\nu d^*_\nu + \ldots = 0 + 0 + \ldots = 0.$$

Remark. In the case that L is *complemented* Theorem 3.1 *is equivalent to the statement:* if $\{a_\lambda\}$ is strongly independent, then the set L_0 of all $\sum x_\lambda$ with $x_\lambda \leqslant a_\lambda$, is a sublattice of L, isomorphic by the correspondence $\sum x_\lambda \leftrightarrow \{x_\lambda\}$ to the direct product of the lattices $[0, a_\lambda]$; L_0 has the property that if *any* family of elements in L_0 has a union or meet in L, then this union or meet is in L_0.

At this point we introduce an important generalization of the conjointness relationship "$(a, b, c)C$" of von Neumann (see § 2).

Definition 3.3. A family of ordered triplets $\{(x_\lambda, c_\lambda, b_\lambda), \lambda \in \Lambda\}$ is called a *C-system* (more precisely, an $\aleph$ *C-system* if $\bar{\Lambda} = \aleph$, and sometimes a *C-sequence* if Λ is countable) if:

(i) $(x_\lambda, c_\lambda, b_\lambda)C$ for every λ, as defined in § 2;

(ii) $\{\sum x_\mu, b_\lambda; \lambda \in \Lambda\}$ is strongly independent.

We shall write: $\{(x_\lambda, c_\lambda, b_\lambda); \lambda \in \Lambda\}C$ to denote that $\{(x_\lambda, c_\lambda, b_\lambda); \lambda \in \Lambda\}$ is a *C*-system.

Clearly, if x denotes $\sum x_\lambda$ and $\{(x_\lambda, c_\lambda, b_\lambda)\}C$, then:

(iii) $x_\lambda = x(c_\lambda + b_\lambda)$ for every λ,

(iv) $x \oplus c_\lambda = x \oplus b_\lambda$ for every λ.

(v) $x \leqslant \sum c_\lambda + \sum b_\lambda$.

Conversely, if some given x, $\{b_\lambda\}$, $\{c_\lambda\}$ satisfy (iv), (v), and

(ii)′ $\{x, b_\lambda; \lambda \in \Lambda\}$ is strongly independent then, with x_λ defined by (iii), it is so that $\{(x_\lambda, c_\lambda, b_\lambda)\}C$ holds.

Thus in a *C*-system the x_λ are uniquely determined by the elements $\{c_\lambda, b_\lambda; \lambda \in \Lambda\}$ and the union $\sum x_\lambda$. We shall sometimes write $\{(x|c_\lambda, b_\lambda)\}C$ with $x = \sum x_\lambda$ in place of $\{(x_\lambda, c_\lambda, b_\lambda)\}C$.

LEMMA 3.1. *If* $\{(x|c_\lambda, b_\lambda)\}C$ *holds, then* $\{c_\lambda\}$ *is residually independent and has residual element* $x\sum c_\lambda$.

Proof.

$$c_\lambda^* \leqslant b_\lambda^* + x.$$

Hence, (by (iv), Definition 3.3),

$$c_\lambda c_\lambda^* \leqslant c_\lambda(b_\lambda + x)(b_\lambda^* + x) = c_\lambda x$$

and (by (iii), Definition 3.3),

$$c_\lambda x = c_\lambda x_\lambda = 0.$$

Thus for each λ, $c_\lambda c^*{}_\lambda = 0$ and hence $\{c_\lambda\}$ is residually independent. Next,

$$x\sum c_\lambda = x(c_\mu^* + c_\mu)(c_\mu^* + x) = x(c_\mu^* + c_\mu(c_\mu^* + x)) = x(c_\mu^* + c_\mu(b_\mu^* + x)).$$

But $\{b_\mu, b^*{}_\mu, x\}$ is independent, so, by the Corollary to Theorem 3.1, $\{c_\mu, b^*{}_\mu, x\}$ is independent; thus $c_\mu(b^*{}_\mu + x) = 0$ and

$$x\sum c_\lambda = x\, c_\mu^* \qquad \text{for each } \mu.$$

Thus $x\sum c_\lambda = x\prod_\mu c^*{}_\mu \leqslant$ (residual element of $\{c_\lambda\}$). On the other hand,

$$\text{(residual element of } \{c_\lambda\}) = \prod c_\lambda^* \leqslant \prod(b_\lambda^* + x) = x$$

since $\{x, b_\lambda; \lambda \in \Lambda\}$ is strongly independent. Thus

$$\text{(residual element of } \{c_\lambda\}) \leqslant (\sum c_\lambda)x$$

and so equality holds.

Definition 3.4. A C-system $\{(x_\lambda, c_\lambda, b_\lambda)\}$ is called a *residual C-system*, if

$$\sum x_\lambda = (\text{residual element of } \{c_\lambda\})$$

equivalently (by Lemma 3.1), if $\sum x_\lambda < \sum c_\lambda$.

Remark. It is easy to see that a residual C-system with Λ *finite* must have all $x_\lambda, c_\lambda, b_\lambda$ identically 0. But a non-trivial residual C-system can be constructed whenever there exists an increasing sequence $\{a_n\}$ which does not converge continuously (this will follow immediately from Theorem 3.6 and the Corollary to Theorem 3.2 below).

THEOREM 3.2. *If $\{a_\lambda\}$ is residually independent and $a_\lambda \geq b_\lambda \oplus c_\lambda$ for every λ, then the residual elements a, b, c (of $\{a_\lambda\}$, $\{b_\lambda\}$, $\{c_\lambda\}$ respectively) satisfy:*

(i) $(\sum b_\lambda)(\sum c_\lambda) = bc$;

(ii) $a \geq b + c$ *with equality if* $a_\lambda = b_\lambda \oplus c_\lambda$ *for every* λ.

Proof. Each of $\{b_\lambda\}$, $\{c_\lambda\}$ is residually independent along with $\{a_\lambda\}$. Now:

(i) For each fixed μ, $\{a^*_\mu, b_\mu, c_\mu\}$ is independent. Hence

$$(\sum b_\lambda)(\sum c_\lambda) = b^*_\mu c^*_\mu; (\sum b_\lambda)(\sum c_\lambda) = (\prod_\mu b^*_\mu)(\prod_\mu c^*_\mu) = bc.$$

(ii) $a \geq b + c$ is clear. But if $a_\lambda = b_\lambda \oplus c_\lambda$ for every λ then also $a \leq b + c$ for:

$$(a + \sum b_\lambda)(\sum c_\lambda) < (b_\mu + a^*_\mu)(\sum c_\lambda) = c^*_\mu.$$

Hence

$$c \geq (a + \sum b_\lambda)(\sum c_\lambda),$$

$$(\sum b_\lambda) + c \geq (a + \sum b_\lambda)(\sum c_\lambda + \sum b_\lambda) \geq a.$$

Similarly $(\sum c_\lambda) + b \geq a$. Thus

$$a \leq ((\sum b_\lambda) + c)((\sum c_\lambda) + b) = b + c + (\sum b_\lambda)(\sum c_\lambda) = b + c.$$

COROLLARY. *If L is complemented and $\{a_\lambda\}$ is residually independent with residual element x, there exists a residual C-system $\{(x|c_\lambda, b_\lambda)\}$ with $c_\lambda \leq a_\lambda$.*

Proof. 1. Choose X to be a complement of x and define:

$$c_\lambda = [a_\lambda - a_\lambda X],$$

$$b_\lambda = (x + c_\lambda)X.$$

2. Then $a_\lambda = a_\lambda X \oplus c_\lambda$ and each of $\{a_\lambda X\}$, $\{c_\lambda\}$ is residually independent since $\{a_\lambda\}$ is residually independent by hypothesis.

But

$$(\text{residual element of } \{a_\lambda X\}) \leq (\text{residual element of } \{a_\lambda\}) = x;$$

also $\leq X$, so $\leq xX = 0$. Now Theorem 3.2 shows that $\{c_\lambda\}$ is residually independent with x as residual element.

3. $\{x, b_\lambda; \lambda \in \Lambda\}$ is strongly independent because

$$b_\lambda^* \leqslant x + c_\lambda^* = c_\lambda^*,$$
$$\prod b_\lambda^* < \prod c_\lambda^* = x.$$

But $\prod b^*_\lambda < X$, hence $\prod b^*_\lambda < xX = 0$. Therefore $\{b_\lambda\}$ is strongly independent.

Since $x\sum b_\lambda < xX = 0$, the Corollary to Theorem 3.1 shows that $\{x, b_\lambda; \lambda \in \Lambda\}$ is strongly independent.

4. $c_\lambda x < a_\lambda x = 0$, $b_\lambda x = 0$ and

$$b_\lambda \oplus x = (x \oplus c_\lambda)(X + x) = x \oplus c_\lambda.$$

Finally, $\sum b_\lambda + \sum c_\lambda > x$ since, as we have already shown, $\{c_\lambda\}$ is residually independent with x as residual element. Thus (iv), (v), and (ii)′ of Definition 3.3. hold and it follows that $\{(x|c_\lambda, b_\lambda)\}$ is a residual C-system, as required.

THEOREM 3.3. *Suppose $\{a_\lambda\}$ is residually independent with residual element t. If $t = x \oplus y$ and Y is an element with $Y > x$ and $Y \oplus y > \sum a_\lambda$, then*

$$\{Y(a_\lambda + y)\}$$

is residually independent with residual element x.

Remark. If L is complemented, Y could be chosen to be $x + [(\sum a_\lambda) - t]$.

Proof. Put $b_\lambda = Y(a_\lambda + y)$. Then

$$b_\lambda + y = a_\lambda + y,\ b_\lambda^* + y = a_\lambda^* + y = a_\lambda^*,$$
$$b_\lambda b_\lambda^* \leqslant Ya_\lambda^*(a_\lambda + y) = Y(a_\lambda^* a_\lambda + y) = Yy = 0.$$

Thus $\{b_\lambda\}$ is residually independent.

Now $b^*_\lambda \leqslant Y$, and $Yy = 0$, so $b^*_\lambda = Y(b^*_\lambda + y) = Ya^*_\lambda$. Hence

$$\text{(the residual element of } \{b_\lambda\}) = \prod(Ya^*_\lambda) = Yt = x + Yy = x.$$

COROLLARY 1. *If L is complemented then every subelement of an $\aleph$-residual element is also an $\aleph$-residual element.*

COROLLARY 2. *If L is complemented and $\{a_\lambda\}$ is residually independent there exists a strongly independent family $\{b_\lambda\}$ with $\sum b_\lambda \leqslant \sum a_\lambda$ and b_λ perspective to a_λ for every λ, with a common axis of perspectivity.*

Proof. Let t be the residual element of $\{a_\lambda\}$ and choose $Y = [(\sum a_\lambda) - t]$, that is, let $x = 0$, $y = t$ in Theorem 3.3. Then $b_\lambda = Y(a_\lambda + y)$ satisfies our requirements and for every λ, b_λ is perspective to a_λ with axis Y.

THEOREM 3.4. *Additivity of perspectivity. Suppose $\{a_\lambda + b_\lambda; \lambda \in \Lambda\}$ is residually independent and $a_\lambda \sim b_\lambda$ for every λ. If $\{a_\lambda\}$ and $\{b_\lambda\}$ are both strongly independent (in particular, if $\{a_\lambda + b_\lambda\}$ is strongly independent),*[10] *then there*

[10]Strong independence of $\{a_\lambda + b_\lambda\}$ obviously implies that of each of $\{a_\lambda\}$, $\{b_\lambda\}$. The interested reader can verify, using Theorem 3.1 and Theorem 3.2, that if L is complemented then residual independence of $\{a_\lambda + b_\lambda\}$ together with strong independence of each of $\{a_\lambda\}$, $\{b_\lambda\}$, actually forces $\{a_\lambda + b_\lambda\}$ to be strongly independent.

exists a perspective mapping of $[0, \sum b_\lambda]$ *onto* $[0, \sum a_\lambda]$ *which maps* b_λ *on* a_λ *for each* λ. *If* $\{b_\lambda\}$ *is strongly independent and* L *is complemented then* $\sum b_\lambda$ *is perspective to a subelement of* $\sum a_\lambda$.

Proof. Suppose $a_\lambda \oplus x_\lambda = b_\lambda \oplus x_\lambda$ and $\{a_\lambda\}$, $\{b_\lambda\}$ are both residually independent. Then for every fixed μ, $(a_\mu + b_\mu)(a^*_\mu + b^*_\mu) = 0$. Hence

$$(\sum x_\lambda)(\sum b_\lambda) = x_\mu b_\mu + x^*_\mu b^*_\mu = 0 + x^*_\mu b^*_\mu \leqslant b^*_\mu;$$

$$(\sum x_\lambda)(\sum b_\lambda) \leqslant (\text{residual element of } \{b_\lambda\}).$$

Similarly,

$$(\sum x_\lambda)(\sum a_\lambda) \leqslant (\text{residual element of } \{a_\lambda\}).$$

Thus if $\{b_\lambda\}$ is strongly independent, $\sum b_\lambda$ is perspective to $[\sum a_\lambda - (\sum x_\lambda)(\sum a_\lambda)]$ with axis $\sum x_\lambda$. If $\{a_\lambda\}$ is also strongly independent then $(\sum x_\lambda)(\sum a_\lambda) = 0$; in this case $(\sum b_\lambda) \sim (\sum a_\lambda)$ with axis $\sum x_\lambda$, and the corresponding perspective mapping maps b_λ on a_λ for each λ.

Remark. If $(\sum b_\lambda)(\sum a_\lambda) = 0$ and $\{b_\lambda\}$ is strongly independent, then residual independence of $\{a_\lambda + b_\lambda\}$ is equivalent to residual independence of $\{a_\lambda\}$, by application of Theorem 3.1.

COROLLARY. *Suppose* $\{a_\lambda\}$, $\{b_\lambda\}$ *are both strongly independent families.*

(i) *If* $a_\lambda \sim b_\lambda$ *for each* λ *and* $(\sum a_\lambda)(\sum b_\lambda) = 0$ *then there is a perspective mapping of* $[0, \sum a_\lambda]$ *onto* $[0, \sum b_\lambda]$ *which maps* a_λ *on* b_λ *for each* λ.

(ii) *If* $a_\lambda \simeq b_\lambda$ *and* L *can be doubled then there is a lattice isomorphism of* $[0, \sum a_\lambda]$ *onto* $[0, \sum b_\lambda]$ *which maps* a_λ *on* b_λ *for each* λ; *if also* $\sum b_\lambda \leqslant \sum a_\lambda$ *and* L *is complemented then* $[\sum a_\lambda - \sum b_\lambda]$ *is a member of an independent sequence of mutually perspective elements.*

Proof of (i): Theorem 3.4 shows this since $\{a_\lambda + b_\lambda\}$ is strongly independent by the Corollary to Theorem 3.1, under the present hypotheses.

Proof of (ii): We may suppose that $L = [0, c]$ with c an element in a modular lattice L_1 such that $[0, c]$ can be mapped by a perspective mapping ϕ onto $[0, u]$ for some u in L_1 with $cu = 0$. Then $a_\lambda \sim \phi(b_\lambda)$ for each λ (by repeated applications of (2.2)).

Since $(\sum a_\lambda)(\sum \phi(b_\lambda)) = 0$, there exists, by (i) above, a perspective mapping ψ of $[0, \sum a_\lambda]$ onto $[0, \sum \phi(b_\lambda)]$ which maps each a_λ on $\phi(b_\lambda)$.

Now $\phi^{-1}\psi$ is a lattice isomorphism of $[0, \sum a_\lambda]$ onto $[0, \sum b_\lambda]$ as required.

If finally $\sum b_\lambda \leqslant \sum a_\lambda$ and L is complemented, let $x_1 = [(\sum a_\lambda) - (\sum b_\lambda)]$; and for $n \geqslant 1$ define x_{n+1} by induction:

$$x_{n+1} = \phi^{-1}\psi(x_n).$$

Then $\{x_n\}$ is an independent sequence since $x_n(\sum (x_m;\ m > n)) = 0$ (this follows from repeated applications of $\phi^{-1}\psi$ to the identity $x_1(\sum (x_m;\ m > 1)) \leqslant x_1(\sum b_\lambda) = 0$).

Since $x_n \sim \phi(x_n)$ and $\phi(x_n) \sim x_{n+1}$ and $\{x_n, \phi(x_n), x_{n+1}\}$ is independent for each n, therefore $x_n \sim x_{n+1}$. Then by (2.2), $x_n \sim x_m$ for all n, m. This proves (ii).

THEOREM 3.5. *Extension of perspective mapping.*

Suppose $\{(x_\lambda, c_\lambda, b_\lambda)\}$ *and* $\{(x_\lambda', c_\lambda', b_\lambda')\}$ *are both C-systems and* $\sum x_\lambda \leqslant x$, $\sum x_\lambda' \leqslant x'$ *and* $\{x + x', \sum b_\lambda, \sum b_\lambda'\}$ *is independent. Then any perspective mapping of* $[0, x]$ *onto* $[0, x']$ *which maps* x_λ *on* x_λ' *for every* λ, *can be extended to a perspective mapping of* $[0, x + \sum b_\lambda]$ *onto* $[0, x' + \sum b_\lambda']$ *which maps* b_λ *on* b_λ' *and* c_λ *on* c_λ' *for every* λ.

Proof. 1 The given perspective mapping of $[0, x]$ onto $[0, x']$ is determined by some axis of perspectivity a with:

$$x \oplus a = x' \oplus a = x + x'.$$

2. We shall choose y_λ below so that

$$y_\lambda \oplus b_\lambda = y_\lambda \oplus b_\lambda' = b_\lambda \oplus b_\lambda'; \tag{3.2}$$

it will then follow immediately, as in the proof of Theorem 3.4, that the axis $a + \sum y_\lambda$ gives a perspective mapping of $[0, x + \sum b_\lambda]$ onto $[0, x' + \sum b_\lambda']$ which fulfills all our requirements except possibly for the requirement:

(3.3) c_λ should be mapped onto c_λ' for each λ.

3. Our choice of y_λ is:

$$y_\lambda = (a + c_\lambda + c_\lambda')(b_\lambda + b_\lambda')$$

and we verify that (3.2) holds, as follows:

(i) $$y_\lambda + b_\lambda = (a + b_\lambda + c_\lambda + c_\lambda')(b_\lambda + b_\lambda')$$

and

$$a + b_\lambda + c_\lambda + c_\lambda' = a + x_\lambda + b_\lambda + c_\lambda' = a + x_\lambda' + b_\lambda + c_\lambda' = a + x_\lambda' + b_\lambda + b_\lambda'$$

so $y_\lambda + b_\lambda = b_\lambda + b_\lambda'$.

Similarly $y_\lambda + b_\lambda' = b_\lambda + b_\lambda'$ so $y_\lambda + b_\lambda = y_\lambda + b_\lambda' = b_\lambda + b_\lambda'$.

(ii) The hypotheses imply that $\{b_\lambda, b_\lambda', x + x'\}$ is an independent family for each λ. Now successive applications of the Corollary to Theorem 3.1 show in turn that each of the following is independent:

$$\{b_\lambda, b_\lambda', x_\lambda', a\}, \qquad \{b_\lambda, c_\lambda', x_\lambda', a\},$$

$$\{b_\lambda, c_\lambda', x_\lambda, a\}, \qquad \{b_\lambda, c_\lambda', c_\lambda, a\}.$$

Therefore

$$b_\lambda y_\lambda = (a + c_\lambda + c_\lambda')b_\lambda = 0.$$

Similarly

$$b_\lambda y_\lambda = 0.$$

(i) and (ii) prove that (3.2) holds.

4. Finally, we verify that $a + \sum y_\lambda$ does satisfy (3.3), as follows:

$$\begin{aligned}(\text{the map of } c_\lambda) &= (a + \textstyle\sum y_\mu + c_\lambda)(x' + \sum b'_\mu) \\ &> (a + y_\lambda + c_\lambda)c'_\lambda = (a + c_\lambda + c'_\lambda)(b_\lambda + b'_\lambda + a + c_\lambda)c'_\lambda \\ &> (b'_\lambda + x_\lambda + a)c'_\lambda > (b'_\lambda + x'_\lambda)c'_\lambda = c'_\lambda,\end{aligned}$$

that is, (map of c_λ) $> c_\lambda'$. Similarly: (map of c_λ') $> c_\lambda$. Since the mappings are inverse perspective mappings, equality must then hold in the preceding two relations and the theorem is completely proved.

THEOREM 3.6. *Suppose L is complemented and $\aleph'$-continuous for every $\aleph' < \aleph$. Suppose also that $\{c_\beta; \beta < \Omega\}$ is an increasing family with $xc_\beta = 0$ for every $\beta < \Omega$ for some fixed x with $x \leqslant \sum c_\beta$. Then there exists a residually independent family $\{a_\beta; \beta < \Omega\}$ such that:*

$$\sum_{\gamma<\beta} a_\gamma = c_\beta \qquad \text{for ever } \beta < \Omega, \tag{3.4}$$

(3.5) *the residual element of $\{a_\beta\}$ is $> x$.*

Proof. By transfinite induction we shall define for each $\beta < \Omega$ a complement C_β of c_β such that $C_\beta > x$ and $C_\gamma > C_\beta$ for all $\gamma < \beta$.

We choose C_1 to be any complement of c_1 with $C_1 > x$. Then for $\beta > 1$, by transfinite induction, we choose C_β to be a relative complement $[\prod_{\delta<\beta} C_\delta - c_\beta(\prod_{\delta<\beta} C_\delta)]$ with $C_\beta > x$. This is possible since, by the inductive assumption, $\prod_{\delta<\beta} C_\delta > x$ and $xc_\beta(\prod_{\delta<\beta} C_\delta) < xc_\beta = 0$; this choice of C_β does give a complement of c_β because

$$\Big(\prod_{\delta<\beta} C_\delta\Big) + c_\beta = \prod_{\delta<\beta} (C_\delta + c_\beta) = \prod_{\delta<\beta} (1) = 1$$

due to the assumption that L is lower $\aleph'$-continuous for $\aleph' < \aleph$.

Now choose $a_1 = c_1$, and for $1 < \beta < \Omega$, choose $a_\beta = c_\beta(\prod_{\delta<\beta} C_\delta)$.

Then (3.4) holds; for by transfinite induction on β, it follows that for every β:

$$\Big(\sum_{\gamma<\beta} c_\gamma + \prod_{\delta<\beta} C_\delta\Big) = \prod_{\delta<\beta} \Big(\sum_{\gamma<\beta} c_\gamma + C_\delta\Big) = \prod_{\delta<\beta} (1) = 1;$$

$$c_\beta = c_\beta \Big(\sum_{\gamma<\beta} c_\gamma + \prod_{\delta<\beta} C_\delta\Big) = \sum_{\gamma<\beta} c_\gamma + c_\beta \Big(\prod_{\delta<\beta} C_\delta\Big) = \sum_{\gamma<\beta} a_\gamma + a_\beta = \sum_{\gamma<\beta} a_\beta.$$

Next, $\{a_\beta; \beta < \Omega\}$ is residually independent; for

$$\begin{aligned}a_\beta(\textstyle\sum(a_\gamma; \gamma \neq \beta)) &= a_\beta c_\beta(\sum(a_\gamma; \gamma \neq \beta)) \\ &= a_\beta(\textstyle\sum(a_\gamma; \gamma < \beta) + c_\beta(\sum(a_\gamma; \gamma > \beta))C_\beta) \\ &= a_\beta(\textstyle\sum(a_\gamma; \gamma < \beta)) = 0\end{aligned}$$

since L is upper $\aleph'$-continuous for $\aleph' < \aleph$.

Finally, for each $\gamma < \Omega$,

$$x = x(\sum c_\beta) \leqslant C_\gamma(\sum a_\beta) = \sum(a_\beta; \beta > \gamma)$$

so (3.5) holds.

COROLLARY 1. *Suppose L is a complemented, $\aleph_0$-complete modular lattice. Then L is upper $\aleph_0$-continuous if and only if every residually independent sequence is strongly independent. More generally, if L is a complemented $\aleph$-complete modular lattice, and L is $\aleph'$-continuous for all $\aleph' < \aleph$, then L is upper $\aleph$-continuous if and only if every $\aleph$-residual element is 0.*[11]

COROLLARY 2. *Suppose L is complemented and $\aleph'$-continuous for every $\aleph' < \aleph$. If $\{a_\lambda\}$ is independent then there exists a strongly independent family $\{b_\lambda\}$ such that $a_\lambda = b_\lambda$ for each λ.*

Proof. We may suppose the a_λ are well-ordered and indexed as $\{a_\beta; \beta < \Omega\}$. Let $c_\beta = \sum(a_\gamma; \gamma \leqslant \beta)$. Then since $\{a_\beta\}$ is independent and L is upper $\aleph'$-continuous for every $\aleph' < \aleph$ it follows that for every $\beta < \Omega$,

$$a_\beta \sum(c_\gamma; \gamma < \beta) = 0, \qquad a_\beta + \sum(c_\gamma; \gamma < \beta) = c_\beta.$$

Now apply Theorem 3.6 (with $x = 0$) to the increasing family $\{c_\beta\}$; it follows that there exists a residually independent family $\{a_\beta'; \beta < \Omega\}$ with

$$\sum(a_\gamma'; \gamma \leqslant \beta) = c_\beta \qquad \text{for every } \beta < \Omega.$$

Clearly for every β, $a_\beta \sim a_\beta'$ with axis $\sum(c_\gamma; \gamma < \beta)$.

Now let the residual element of $\{a_\beta'\}$ be denoted as y and let Y be a complement of y. Then by Theorem 3.3 the elements $b_\beta = Y(a_\beta' + y)$ form a strongly independent family.

Since $b_\beta \sim a_\beta'$ (with axis y) for each β, and $a_\beta' \sim a_\beta$ (as shown above), therefore $b_\beta = a_\beta$ and so the b_β satisfy our requirements.

4. Additivity of continuity. In this section and in §§ 5, 6, 7, we assume that L is a complemented, $\aleph$-complete modular lattice.

THEOREM 4.1. *Suppose $a_\lambda \uparrow a$, $b_\lambda \uparrow b$, and both a_λ, b_λ converge continuously. If $a_\lambda b_\lambda = 0$ for every λ (equivalently, if $ab = 0$), then $a_\lambda + b_\lambda$ also converges continuously.*

Proof. We may suppose $x(a_\lambda + b_\lambda) = 0$ for every λ (which implies $x(a_\lambda + b_\mu) = 0$, $(x + a_\lambda)b_\mu = 0$ for all λ, μ) and need only prove that $x(a + b) = 0$. But the continuous convergence of b_μ yields for every λ, $(x + a_\lambda)b = 0$, and so $(x + b)a_\lambda = 0$; continuous convergence of a_λ yields $(x + b)a = 0$, hence

$$x(a + b) = x(a(x + b) + b) = xb = \sum(xb_\lambda) = 0 \text{ as required.}$$

THEOREM 4.2. *If $[a, 1]$ is upper $\aleph$-continuous and $c_\lambda \uparrow 1$ in L then $\sum(ac_\lambda) = a$.*

Proof. 1. First consider the case that for every λ, $ac_\lambda = 0$. We shall show

[11] In the case $\aleph = \aleph_0$ *no* continuity assumption is implied.

that in this case $(a, c_\mu)P$ holds for each μ. Then since $\sum c_\mu = 1$, this implies $(a, 1)P$, hence $a = 0$, as required.

To show $(a, c_\mu)P$ holds we let C_μ be an arbitrary complement of c_μ and we need only prove that $C_\mu \geqslant a$ (see 2.6)). But if $\lambda \geqslant \mu$,

$$c_\lambda = c_\lambda(c_\mu \oplus C_\mu) = c_\mu + c_\lambda C_\mu.$$

Hence

$$\begin{aligned} c_\mu \oplus C_\mu = 1 = \textstyle\sum_\lambda c_\lambda &= \textstyle\sum_\lambda (c_\mu + c_\lambda C_\mu) \\ &= c_\mu \oplus \textstyle\sum_\lambda (c_\lambda C_\mu) \end{aligned}$$

by the definition of lattice union. Since $C_\mu \geqslant \sum_\lambda (c_\lambda C_\mu)$, the modular law implies that $C_\mu = \sum_\lambda (c_\lambda C_\mu)$. Hence $a + C_\mu = a + \sum_\lambda (c_\lambda C_\mu) = \sum_\lambda (a + c_\lambda C_\mu)$, by the definition of lattice union. Then

$$\begin{aligned} (a + C_\mu)(a + c_\mu) &= (\textstyle\sum_\lambda (a + c_\lambda C_\mu))(a + c_\mu) \\ &= \textstyle\sum_\lambda ((a + c_\lambda C_\mu)(a + c_\mu)) \end{aligned}$$

since $[a, 1]$ is upper $\aleph$-continuous,

$$\begin{aligned} &= \textstyle\sum_\lambda (a + c_\mu (a + c_\lambda C_\mu)) \\ &= a \end{aligned}$$

because

$$c_\mu(a + c_\lambda C_\mu) = c_\mu(a(c_\mu + c_\lambda C_\mu) + c_\lambda C_\mu) \leqslant c_\mu(ac_\lambda + C_\mu) = c_\mu(0 + C_\mu) = 0.$$

Thus, in turn,

$$\begin{aligned} (a + C_\mu)c_\mu = (a + C_\mu)(a + c_\mu)c_\mu = ac_\mu = 0; \\ (a + C_\mu)c_\mu + C_\mu = C_\mu; \\ a + C_\mu = C_\mu; \\ a \leqslant C_\mu \end{aligned}$$

as required.

2. In the general case, let $a_0 = \sum_\lambda (ac_\lambda)$. Then $a_0 \leqslant a$ and $(a_0 + c_\lambda)a = a_0$ for every λ. Since $(a_0 + c_\lambda) \uparrow 1$ in the lattice $[a_0, 1]$, we can apply the argument of the preceding paragraph with $[a_0, 1]$ in place of L. We obtain: $a = a_0$, that is, $\sum (ac_\lambda) = a$, as required.

THEOREM 4.3. *Additivity of upper $\aleph$-continuity. If both $[0, a]$, $[0, b]$ are upper $\aleph$-continuous then $[0, a + b]$ is also upper $\aleph$-continuous.*

Proof. We may clearly suppose $a \oplus b = 1$, $c_\lambda \uparrow 1$ and need only prove $(xc_\lambda) \uparrow x$ for every x. But ac_λ, bc_λ both converge continuously; hence, by Theorem 4.1, $ac_\lambda + bc_\lambda$ converges continuously.

By (2.1), $[a, 1]$ is lattice isomorphic to $[0, b]$ and hence is upper $\aleph$-continuous. Then, by Theorem 4.2, $\sum (ac_\lambda) = a$. Similarly $\sum (bc_\lambda) = b$. So $\sum (ac_\lambda + bc_\lambda) = a + b = 1$.

But we have shown that $ac_\lambda + bc_\lambda$ converges continuously; so for every x, $x \geqslant \sum(xc_\lambda) \geqslant \sum x(ac_\lambda + bc_\lambda) = x$. This shows that $(xc_\lambda) \uparrow x$ and proves Theorem 4.3.

THEOREM 4.4. (Generalization of Theorem 4.3.) *Suppose L is upper $\aleph'$-continuous for some $\aleph' < \aleph$ and $[0, a_\mu]$ is upper $\aleph$-continuous for each $\mu \in \Gamma$ with $\Gamma \leqslant \aleph'$. Then $[0, \sum a_\mu]$ is upper $\aleph$-continuous.*

Proof. 1. We may suppose that $\aleph'$ is infinite since Theorem 4.3 shows that Theorem 4.4 holds for finite $\aleph'$.

2. We shall prove Theorem 4.4 by transfinite induction on $\aleph'$; we may therefore suppose that Theorem 4.4 holds for all cardinals less than the given infinite $\aleph'$.

3. We may now suppose that the indices μ are arranged as the set of ordinal numbers $\beta < \Omega_1$, where Ω_1 is the least ordinal number of corresponding cardinal power $\aleph'$.

4. Since $[0, \sum(a_\beta; \beta < \gamma)]$ is upper $\aleph$-continuous for every $\gamma < \Omega_1$ (by the inductive assumption), we may assume that $\{a_\beta\}$ is increasing, say $a_\beta \uparrow a$.

5. Thus we may suppose:

(i) For each $\beta < \Omega_1$, $[0, a_\beta]$ is upper $\aleph$-continuous and $a_\beta \uparrow a$ with continuous convergence (since L is assumed to be upper $\aleph'$-continuous).

And we need only prove that $[0, a]$ is upper $\aleph$-continuous.

It is sufficient to prove:

(ii) $c_\gamma \uparrow a$, $xc_\gamma = 0$ for all $\gamma < \Omega_2$ for some $\Omega_2 \leqslant \Omega$ together imply $xa = 0$.

6. For each β, $(c_\gamma a_\beta) \uparrow \bar{a}_\beta$ where $\bar{a}_\beta = \sum_\gamma (c_\gamma a_\beta) \leqslant a_\beta$.

Clearly $\{\bar{a}_\beta; \beta < \Omega_1\}$ is an increasing family, along with $\{a_\beta\}$, and converges continuously since L is assumed to be upper $\aleph'$-continuous. Hence, for every γ,

$$c_\gamma(\textstyle\sum_\beta \bar{a}_\beta) = \sum_\beta (c_\gamma \bar{a}_\beta).$$

Now

$$\begin{aligned}
\textstyle\sum \bar{a}_\beta &= \textstyle\sum_{\gamma,\beta}(c_\gamma a_\beta) = \sum_\gamma(\sum_\beta c_\gamma a_\beta) \\
&= \textstyle\sum_\gamma(c_\gamma a) \quad \text{since } a_\beta \uparrow a, \text{ continuous convergence,} \\
&= \textstyle\sum_\gamma c_\gamma \quad \text{since } c_\gamma \leqslant a \text{ for every } \gamma, \\
&= a \quad \text{since } c_\gamma \uparrow a, \text{ by hypothesis.}
\end{aligned}$$

Thus $\bar{a}_\beta \uparrow a$ and the convergence is continuous.

Next, for each β,

$$\begin{aligned}
x\bar{a}_\beta &= x\textstyle\sum_\gamma(c_\gamma a_\beta) \\
&= \textstyle\sum_\gamma(xc_\gamma a_\beta) \quad \text{since } [0, a_\beta] \text{ is upper } \aleph\text{-continuous,} \\
&= \textstyle\sum_\gamma(0) \quad \text{since } xc_\gamma = 0 \text{ for very } \gamma, \\
&= 0.
\end{aligned}$$

This proves the theorem.

5. Homogeneous sequences. We assume, as in § 4, that L is a complemented $\aleph$-complete, modular lattice but most of this section involves only the complemented countably complete modular lattices.

Definition 5.1. A sequence $\{a_n\}$ is called *homogeneous* if $\{a_n\}$ is strongly independent and the a_n are pairwise perspective.

Definition 5.2. If $\{a_n\}$ is a sequence in L then for any complement A of $\sum a_n$, the sequence $\{a^*_n + A\}$ is called a *dual sequence* of $\{a_n\}$.

Remark 1. Each dual sequence of $\{a_n\}$ is strongly independent in L'; if $\{a_n\}$ is strongly independent in L then each of its dual sequences, considered in L', has the original $\{a_n\}$ as a dual sequence.

Remark 2. If $\{a_n\}$ is homogeneous in L then each of its dual sequences is homogeneous in L'.

Definition 5.3. A homogeneous sequence $\{a_n\}$ is said to be of type (A) if all the a_n possess a common complement (equivalently, a common relative complement in $\sum a_n$), that is, there exists an element A such that $a_n \oplus A = 1$ for all n.

A homogeneous sequence $\{a_n\}$ is said to be of type (A*) if all the a^*_n have a common complement (equivalently, a common relative complement in $\sum a_n$).

Remark 1. Clearly if $\{a_n\}$ is strongly independent, then $\{a_n\}$ is homogeneous and of type (A), or (A*), if and only if one (hence all) of its dual sequences is homogeneous and of type (A*), or (A) respectively, in L'.

Hence, if every homogeneous sequence in L is of type (A), or if every homogeneous sequence in L is of type (A*), then every homogeneous sequence in L' is of type (A*) or (A), respectively.

Remark 2. If $\{a_n\}$ is a homogeneous sequence and $x_1 \leqslant a_1$, then any set of perspective mappings of $[0, a_1]$ onto $[0, a_n]$ when applied to x_1 will yield a homogeneous sequence $\{x_n\}$ (Theorem 5.1 below and its Corollary 1 will imply that if $\{a_n\}$ is of type (A), or (A*), then $\{x_n\}$ has the same property).

Remark 3. If $\{a_n\}$ is a homogeneous sequence then every infinite subsequence is also homogeneous; and if $\{a_n\}$ is of type (A), or (A*), then every infinite subsequence is of the same type.

LEMMA 5.1. *If* $\{a_0, a_1, \ldots, a_n, \ldots\}$ *is a homogeneous sequence of type* (A) *there exists at least one C-sequence* $\{(a_0|c_n, a_n)\}$ *such that* $a_0 \sum c_n = 0$.

Remark. From Theorem 5.1 below it will follow that under the hypothesis of Lemma 5.1 *every C*-sequence $\{(a_0|c_n, a_n)\}$ has the property $a_0 \sum c_n = 0$.

Proof. Let A be a common complement of the a_n. Choose $c_n = A(a_0 + a_n)$ for $n \geqslant 1$. Then

$$a_0 c_n = a_n c_n = 0, \qquad a_0 \oplus c_n = a_0 \oplus a_n = c_n \oplus a_n.$$

The lemma follows since $\{a_0, a_1, \ldots\}$ is strongly independent (by hypothesis) and $a_0 \sum c_n \leqslant a_0 A = 0$.

LEMMA 5.2. *Suppose* $\{a_0, a_1, \ldots\}$ *is a homogeneous sequence of type* (A) *and* $A \sum a_n = 0$. *If* $x \leqslant A$, $(x, \sum a_n)P$ *together imply* $x = 0$ (*in particular, if* A *is perspective to a subelement of* $\sum a_n$), *then* $[0, A]$ *is upper* $\aleph_0$*-continuous.*

Proof 1. By Corollary 1 to Theorem 3.6 we need only prove that every residually independent sequence in $[0, A]$ has residual element zero. We may therefore suppose that x_1 $(\neq 0)$ is the residual element of some residually independent sequence in $[0, A]$ and we need only derive a contradiction.

2. The hypotheses imply that $(x_1, \sum a_n)P$ is false; therefore $(x_1, a_n)P$ is false for some n, hence $(x_1, a_0)P$ is false since $a_n \sim a_0$. Thus there exists $x \neq 0$ with $x \leqslant x_1$ and x perspective to a subelement of a_0. Theorem 3.3 shows that x is the residual element of some residually independent sequence in $[0, A]$.

By Remark 2 following Definition 5.3 we may suppose (by replacement of a_n by suitable subelements) that x is perspective to a_0, say by a perspective mapping ϕ.

3. By the Corollary to Theorem 3.2 there exists a residual C-sequence $\{(x|c_n, b_n)\}$ with $\sum c_n + \sum b_n \leqslant A$; then x is the residual element of $\{c_n\}$ and $x = \sum x_n$ for suitable x_n such that $\{(x_n, c_n, b_n)\}C$ holds.

4. By Lemma 5.1 there exists a C-sequence $\{(a_0|d_n, a_n)\}$ with $a_0 \sum d_n = 0$. We shall derive a contradiction in the following way: we shall construct a C-sequence $\{(x_n', c_n', b_n')\}$ with:

(i) $$x_n' = \phi(x_n)$$

(ii) $$c_n' \leqslant d_n, b_n' \leqslant a_n.$$

(i) will imply that $\sum x_n' = \phi(\sum x_n) = \phi(x) = a_0$ and (ii) will imply that $(\sum c_n')(\sum x_n') = 0$. Then the "extension of perspective mapping" Theorem 3.5 will apply and give an extension of ϕ (which we write again as ϕ) such that $\phi(c_n) = c_n'$ for all n. This will yield:

$$\begin{aligned} \phi(x) &= \phi(x \textstyle\sum c_n) = \phi(x) \textstyle\sum \phi(c_n) \\ &= a_0 \textstyle\sum c_n' \leqslant a_0 \sum d_n = 0 \end{aligned}$$

and imply that $x = 0$, the desired contradiction.

5. The reader can verify easily that the elements $x_n' = \phi(x_n)$, $c_n' = (x_n' + a_n)d_n$, $b_n' = (a_0 + c_n')a_n$ form a C-sequence satisfying (i), (ii) above. Thus Lemma 5.2 is proved.

LEMMA 5.3. *Suppose* $\{x, a_n; n > 1\}$ *and* $\{x, b_n; n > 1\}$ *are homogeneous sequences with* $\{x, \sum a_n, \sum b_n\}$ *independent. Then* $S = \{x, a_1, b_1, a_2, b_2, \ldots\}$ *is a homogeneous sequence. Moreover if both* $\{x, a_n; n > 1\}$ *and* $\{x, b_n; n > 1\}$ *are of type* (A) *then* S *is also of type* (A).

Proof. S is strongly independent, by application of the Corollary to Theorem 3.1. Moreover, all of $x, a_n (n > 1), b_n (n > 1)$ are pairwise perspective so S is a homogeneous sequence.

Now suppose A and B are common relative complements of each of x, a_n in $x + \sum a_n$ and of each of x, b_n in $x + \sum b_n$, respectively.

Then $A + B$ is a common relative complement of each of x, a_n, b_n in $x + \sum a_n + \sum b_n$; for if c is x or a_p then

$$c(A + B) = cA = 0 \quad \text{since} \quad B(x + \sum a_n)(x + \sum b_n) = Bx = 0.$$

Similarly $c(A + B) = 0$ if c is b_p. It is clear that $c + A + B = x + \sum a_n + \sum b_n$ if c is x, a_p, or b_p. This proves Lemma 5.3.

Now we shall prove:

THEOREM 5.1. *The following conditions are equivalent for a homogeneous sequence* $\{a_n\}$:

(i) $[0, \sum a_n]$ *is upper* $\aleph_0$*-continuous.*

(ii) $\sum_{i=1}^{n} a_i$ *converges continuously.*

(iii) $\{a_n\}$ *is of type* (A).

Proof. 1. (i) *implies* (ii): this is trivial.

2. (ii) *implies* (iii): By Lemma 5.3 it is sufficient to prove that $\{a_{2n}\}$ is of type (A).

Suppose $(a_{2n}, x_n, a_{2n+2})C$. Then $\{x_n\}$ is a homogeneous sequence with $x_n \sim a_{2n-1}$ for all $n > 1$ (use (3.1) and (2.2)). Since $(\sum x_n)(\sum a_{2n-1}) \leqslant (\sum a_{2n})(\sum a_{2n-1}) = 0$, Theorem 3.4 shows that there exists a perspective mapping of $[0, \sum x_n]$ onto $[0, \sum a_{2n-1}]$ which maps x_n on a_{2n-1}.

But $\sum_{i=1}^{n} a_{2i-1}$ converges continuously: to see this, observe that for every y,

$$y \sum a_{2i-1} = \Big(y \sum a_{2i-1}\Big)\Big(\sum a_i\Big) = \sum_{n=1}^{\infty} \Big(y \sum a_{2i-1}\Big)\Big(\sum_{j=1}^{2n} a_j\Big)$$

$$= \sum_{n=1}^{\infty} \Big(y \sum_{i=1}^{n} a_{2i-1}\Big).$$

Therefore, $\sum_{i=1}^{n} x_i$ converges continuously. But $a_{2p} \sum_{i=1}^{n} x_i = 0$ for every n, so $x = \sum x_i$ satisfies $a_{2p} x = 0$ for every p. Obviously $a_{2p} + x = \sum a_{2n}$ for every p so the a_{2p} all have x as common relative complement in $\sum a_{2n}$. This proves that $\{a_{2n}\}$ is of type (A) and shows that (ii) implies (iii).

3. (iii) *implies* (i): In Lemma 5.1 use $A = \sum a_{2n}$. Since $\{a_{2n-1}\}$ is of type (A) and $A \sum a_{2n-1} = 0$, $A \sim \sum a_{2n-1}$, therefore Lemma 5.1 applies and shows that $[0, \sum a_{2n}]$ is upper $\aleph_0$-continuous. Similarly $[0, \sum a_{2n-1}]$ is upper $\aleph_0$-continuous.

Now by Theorem 4.3 (the additivity of continuity), $[0, \sum a_n]$ is upper $\aleph_0$-continuous.

COROLLARY 1. *The following are equivalent for a homogeneous sequence* $\{a_n\}$:

(i) $[0, \sum a_n]$ *is lower* $\aleph_0$*-continuous.*

(ii) $\sum_{i=n}^{\infty} a_i$ *converges continuously.*

(iii) $\{a_n\}$ *is of type* (A*).

Proof. Apply Theorem 5.1 to $\{a^*_n\}$ in the lattice dual to $[0, \sum a_n]$.

COROLLARY 2. *If a homogeneous sequence* $\{a_n\}$ *is of type* (A) *and also of type* (A*) *then all* $a_n = 0$.

Proof. Let A be a common relative complement of the a^*_m in $\sum a_n$. Since $A\sum_{i=1}^{m} a_i \leqslant Aa^*_{m+1} = 0$, and $\sum_{i=1}^{m} a_i$ converges continuously by Theorem 5.1, therefore $A = 0$. Then $a^*_m = \sum a_n$, $a_m \leqslant a_m a^*_m = 0$, so all a_m are 0.

COROLLARY 3. *If* $\{a_n\}$ *is a homogeneous sequence of type* (A) *and* $[0, a_n]$ *is upper* $\aleph$*-continuous for every* n, *then* $[0, \sum a_n]$ *is also upper* $\aleph$*-continuous.*

Proof. This follows from Theorem 4.4.

Remark. Corollary 3 applies, in particular, if each a_n is an atom.

6. Additivity of finiteness. We assume, as in §§ 4, 5, that L is a complemented $\aleph$-complete modular lattice.

Definition 6.1. L is called *finite* if every independent sequence of pairwise perspective elements has all its elements zero.[12]

THEOREM 6.1. *If* $\{c_n\}$ *is an independent sequence of pairwise perspective elements there exists a homogeneous sequence* $\{d_n\}$ *with* $d_1 = c_1$, $d_m \sim c_n$ *for all* m, n *and* $\sum d_n \leqslant \sum c_n$.

Proof. By Theorem 3.6, applied to our $\sum_{i=1}^{n} c_i$ with $x = 0$ (no continuity is required in the hypotheses for the case $\aleph = \aleph_0$), there exists a residually independent sequence $\{a_n\}$ with

[12]The following possible definitions of "finiteness" for a modular lattice with zero:

(F_1): as in Definition 6.1,
(F_2): $a = b$, $b \leqslant a$, imply $a = b$,
(F_3): $a \sim c$, $c \sim b$, $b < a$ imply $a = b$

are related as follows:

(i) (F_2) implies (F_3) always.

(ii) (F_1) implies (F_2) if the lattice is also complemented.

(iii) If the lattice is also complemented and countably complete then (F_1), (F_2), and (F_3) are all equivalent.

(iv) If the lattice is not countably complete then (F_2) need not imply (F_1); this is shown by the example of footnote 14 where the lattice is even orthocomplemented and perspectivity is actually transitive.

(i) is trivially true.

To prove (ii): suppose there is a projective mapping ϕ of $[0, a]$ onto $[0, b]$ with $b < a$ and $b \neq a$. The argument used to prove (2.7) actually shows that for some $0 \neq x_1 \leqslant [a - b]$ we have $x_1 \sim x_2$ where $x_2 = \phi(x_1)$. Let $x_n = \phi(x_{n-1})$ for $n > 1$. Then repeated application of ϕ to the relation $x_1 \sim x_2$ shows that $\{x_n\}$ is independent and pairwise perspective, so (F_1) fails to hold. This proves (ii).

To prove (iii): suppose $\{a_n\}$ pairwise perspective and independent. By Theorem 6.1, with the same a_1, we may assume even strong independence. Then by Theorem 3.4,

$$\Sigma(a_{2n-1}; n > 1) \sim \Sigma(a_{2n}; n > 1), \Sigma(a_{2n}; n > 1) \sim \Sigma(a_{2n+1}; n > 1).$$

Now (F_2) would force $a_1 = 0$, so (F_3) implies (F_1) and (iii) follows from the previous remarks.

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} c_i$$

for every $n \geqslant 1$. We observe that

$$a_1 = c_1 \text{ and for } n > 1, a_n \sim c_n \left(\text{axis } \sum_{i=1}^{n-1} a_i = \sum_{i=1}^{n-1} c_i \right).$$

Now the a_n are pairwise perspective; for if $p > n$ then $c_p(a_n + c_n) = 0$ so (2.2) implies $a_n \sim c_p$. If $m \neq n$ and $p \geqslant m$ and $p \geqslant n$, then $\{a_m, a_n, c_p\}$ is independent, $a_m \sim c_p \sim a_n$, so by (2.2), $a_m \sim a_n$.

If $p < n$ then $a_n(a_p + c_p) = 0$ so $a_p \sim a_n$, $a_p \sim c_p$ yield by (2.2) that $a_n \sim c_p$. Thus $a_n \sim c_p$ for all n, p and $a_1 = c_1$.

Now let y be the residual element of $\{a_n\}$ and let Y be a relative complement of y in $\sum a_n = \sum c_n$ with $Y \geqslant a_1 = c_1$. Let $d_n = (y + a_n)Y$.

Then Theorem 3.3 shows that $\{d_n\}$ is strongly independent. Now d_n is the map of a_n in a perspective mapping of

$$\left[0, \sum_{i=1}^{n} a_i\right] \text{ onto } \left[0, \sum_{i=1}^{n} d_i\right]$$

with axis y, for any $m \geqslant n$. Hence the d_n are pairwise perspective, along with the a_n. Thus $\{d_n\}$ is a homogeneous sequence.

The definitions show that $d_1 = a_1 = c_1$ and for every n,

$$y + \sum_{i=1}^{n} d_i = y + \sum_{i=1}^{n} a_i = y + \sum_{i=1}^{n} c_i.$$

If $n > 1$, then $d_n \sim c_n$ with axis $y + \sum_{i=1}^{n-1} c_i$ (use: $y \sum_{i=1}^{n} c_i = y \sum_{i=1}^{n} a_i = 0$ and $y \sum_{i=1}^{n} d_i \leqslant yY = 0$).

But then $d_m \sim c_n$ for all $m \neq n$; for $d_m \sim d_1 = c_1 \sim c_n$ and $\{d_m, c_1, c_n\}$ is independent, so (2.2) yields $d_m \sim c_n$.

Since $\sum d_n \leqslant \sum c_n$ obviously, Theorem 6.1 is proved.

COROLLARY. *If every homogeneous sequence has all its elements zero then the lattice is finite.*

LEMMA 6.1. *Suppose L is $\aleph'$-continuous for all $\aleph' < \aleph$ and suppose L can be doubled. If $\{a_\beta; \beta < \Omega\}$ is strongly independent and $x \leqslant \sum a_\beta$ but $x \sum (a_\gamma; \gamma \leqslant \beta) = 0$ for all $\beta < \Omega$ then x is a member of a homogeneous sequence.*

Proof. Let $X = [(\sum a_\beta) - x]$ and define $\bar{a}_\beta = (a_\beta + x)X$. Then $\{\bar{a}_\beta\}$ is obviously an independent family and $a_\beta \sim \bar{a}_\beta$ for each β. Now Corollary 2 to Theorem 3.6, applied to $[0, X]$, gives a strongly independent family $\{b_\beta\}$ with $b_\beta \leqslant X$ and $\bar{a}_\beta = b_\beta$.

Since $\{a_\beta\}$, $\{b_\beta\}$ are both strongly independent and $a_\beta = b_\beta$ for every β, and $\sum b_\beta \leqslant \sum a_\beta$, therefore (ii) of the Corollary to Theorem 3.4, together with Theorem 6.1, show that $[\sum a_\beta - \sum b_\beta]$ is a member of a homogeneous

sequence. But the relative complement $[\sum a_\beta - \sum b_\beta]$ could be chosen $\geqslant x$ so (use Remark 2 following Definition 5.3) x itself is a member of a homogeneous sequence, as stated.

THEOREM 6.2. *The following properties are equivalent:*

(i) *every homogeneous sequence in L is of type* (A),

(ii) *for every strongly independent sequence $\{a_n\}$ for which $[0, \sum a_n\}$ can be doubled, $\sum_{i=1}^{n} a_i$ converges continuously.*

Proof. 1. (ii) implies (i): Let $\{x_n\}$ be a homogeneous sequence. Then each of $\{x_{2n}\}$, $\{x_{2n-1}\}$ is strongly independent (in fact, a homogeneous sequence) and each of $\sum x_{2n}$, $\sum x_{2n-1}$ can be doubled (in fact, $(\sum x_{2n})(\sum x_{2n-1}) = 0$ and $(\sum x_{2n}) \sim (\sum x_{2n-1})$ by (i) of the Corollary to Theorem 3.4).

Now if (ii) holds, then each of

$$\sum_{i=1}^{n} x_{2i}, \quad \sum_{i=1}^{n} x_{2i-1}$$

converges continuously so by Theorem 5.1, each of $[0, \sum x_{2n}]$, $[0, \sum x_{2n-1}]$ is upper $\aleph_0$-continuous; hence $[0, \sum x_n]$ is upper $\aleph_0$-continuous by Theorem 4.3. Finally $\{x_n\}$ is of type (A) by Theorem 5.1. So (ii) implies (i).

2. (i) *implies* (ii). Suppose (i) holds. We may suppose that $\{a_n\}$ is strongly independent and that the lattice $[0, \sum a_n]$ can be doubled and we need only prove that $\sum_{i=1}^{n} a_i$ converges continuously.

We may suppose that there exists an element $x \neq 0$ such that $x \leqslant \sum a_n$ and $x \sum_{i=1}^{n} a_i = 0$ for all n and we need only derive a contradiction.

3. By replacing each a_n by $a_n(x + a^*_n)$ we may even suppose that a_n is perspective to a subelement of x (observe: $\sum a_n(x + a^*_n) \leqslant \sum a_n$, so $[0, \sum a_n(x + a^*_n)]$ can be doubled and has property (i) along with $[0, \sum a_n]$; also $\{a_n(x + a^*_n)\}$ is strongly independent, along with $\{a_n\}$; finally, $a_n(x + a^*_n) \sim [x - xa^*_n]$ with axis a^*_n).

4. We shall show now that $[0, \sum a_n]$ is upper $\aleph_0$-continuous; this implies that $x = 0$ and gives the desired contradiction.

5. In the present situation, Lemma 6.1 applies, with $[0, \sum a_n]$ in place of L, and shows that there exists a homogeneous sequence $\{x_n\}$ in $[0, \sum a_n]$ with $x = x_1$. The validity of (i) in $[0, \sum a_n]$ then implies that $\{x_n\}$ is of type (A).

6. Since $[0, \sum a_n]$ can be doubled we may (and shall) assume that $[0, \sum a_n]$ is identified with $[0, v]$ in some modular lattice with zero, L_1, in such a way that there exists a perspective mapping ϕ of $[0, \sum a_n]$ onto $[0, u]$ for some u in L_1 with $u \sum a_n = 0$ (we do not know that L_1 is complemented and has property (i) but $[0, \sum a_n]$, but so also $[0, u]$, does have these properties). Then $\{\phi(x_n)\}$ is a homogeneous sequence of type (A) along with $\{x_n\}$.

Since a_n is perspective to a subelement of x, and x is perspective to x_n and $(a_n + x + x_n)\phi(x_n) = 0$, (2.2) and $x_n \sim \phi(x_n)$ imply that a_n is perspective to a subelement of $\phi(x_n)$.

By (i) of the Corollary to Theorem 3.4, $\sum a_n$ is perspective to a subelement

of $\sum\phi(x_n)$. But $[0,\sum\phi(x_n)]$ is upper $\aleph_0$-continuous by Theorem 5.1, hence $[0,\sum a_n]$ is also upper $\aleph_0$-continuous. This completes the proof of Theorem 6.2.

LEMMA 6.2. *If $\{c_n\}$ is a strongly independent sequence and a is an arbitrary element, there exists a decomposition $c_n = c_n' \oplus c_n''$ with the properties:*

(i) *$c_n' \sim d_n$ for some strongly independent $\{d_n\}$ with $\sum d_n \leqslant a(\sum c_n)$,*

(ii) *$a\sum c_n'' = 0$.*

Proof. 1. Put

$$c_n' = c_n\Big(a + \sum_{m>n} c_m\Big), \qquad c_n'' = [c_n - c_n'].$$

2. (ii) is immediate since

$$a\Big(\sum_{m=n}^{\infty} c_m''\Big) = a\Big(c_n''\Big(a + \sum_{m>n} c_m''\Big) + \sum_{m>n} c_m''\Big)$$
$$= a\sum_{m>n} c_m'';$$
$$a(\sum c_m'') = a\sum(c_m''; m > 2) = \ldots = a\sum(c_m''; m > n)$$
$$= a\prod_n \sum(c_m''; m > n) \leqslant \prod_n c_n^* = 0.$$

3. To prove (i) we note that

$$c_n' \oplus \sum_{m>n} c_m \leqslant a + \sum_{m>n} c_m$$

hence

$$c_n' \oplus \sum_{m>n} c_m = d_n \oplus \sum_{m>n} c_m \tag{6.1}$$

for suitable $d_n \leqslant a$.

Now (6.1) shows that $d_n \sim c_n'$ (axis $\sum_{m>n} c_m$) and $\{d_n\}$ is strongly independent by (3.1).

Since each $d_m \leqslant \sum c_n$, and $\leqslant a$, therefore $\sum d_n \leqslant a\sum c_n$ and Lemma 6.2 is proved.

THEOREM 6.3. *If in $[0, a]$ and in $[0, b]$ every homogeneous sequence is of type* (A) *then this is true in $[0, a + b]$.*

Proof. We may suppose $a \oplus b = 1$. By Theorem 6.2 we need only prove: if $\{c_n\}$ is strongly independent and $[0,\sum c_n]$ can be doubled, then $\sum_{i=1}^{n} c_i$ converges continuously.

We shall use the decomposition of c_n, $c_n = c_n' \oplus c_n''$ and the d_n, provided by Lemma 6.2 for the present a. We shall show:

(i) $\sum_{i=1}^{n} c_i'$ converges continuously,

(ii) $\sum_{i=1}^{n} c_i''$ converges continuously.

It will then follow from Theorem 4.1 that $\sum_1^n c_n$ converges continuously, proving the theorem.

To prove (ii): We note that $[0, \sum c_n'']$ is mapped by a perspective mapping (axis a) on a sublattice of $[0, b]$. Since this sublattice can be doubled (it is lattice isomorphic to $[0, \sum c_n'']$ and $\sum c_n'' \leqslant \sum c_n$), (ii) follows from the assumed properties of $[0, b]$.

To prove (i), we observe that (ii) of the Corollary to Theorem 3.4 applies to the lattice $[0, \sum c_n]$ since $\{d_n\}$, $\{c_n'\}$ are each strongly independent, $d_n \sim c_n''$ for each n, $d_n \leqslant \sum c_m$, $c_n'' \leqslant \sum c_m$ and $[0, \sum c_n]$ can be doubled. Thus $[0, \sum c_n'']$ is lattice isomorphic to $[0, \sum d_n]$.

Since $[0, \sum d_n]$ can be doubled (along with $[0, \sum c_n'']$, along with $[0, \sum c_n]$) and since $\sum d_n \leqslant a$, it follows from the hypothesis that $\sum_{i=1}^{n} d_i$ converges continuously. Hence $\sum_{i=1}^{n} c_i''$ also converges continuously.

This proves (i) and completes the proof of Theorem 6.3.

COROLLARY 1. *If in* $[0, a]$ *and in* $[0, b]$ *every homogeneous sequence is of type* (A*) *then this is also true in* $[0, a + b]$.

Proof. We may suppose $a \oplus b = 1$. Now Theorem 6.3 (for L') implies: if in each of $[a, 1]'$, $[b, 1]'$ *considered as sublattices of* L', every homogeneous sequence is of type (A), then this is true in $[ab, 1]'$, that is, $[0, 1]'$.

But $[a, 1]'$, $[b, 1]'$ are anti-isomorphic to $[0, b]$, $[0, a]$ respectively, by (2.1). Thus, if every homogeneous sequence in $[0, a]$ or $[0, b]$ is of type (A*) then every homogeneous sequence in $[b, 1]'$ or $[a, 1]'$ is of type (A) (use the Remark 1 following Definition 5.3); hence every homogeneous sequence in $[0, 1]'$ is of type (A); finally every homogeneous sequence in $[0, 1]$ is of type (A*) (again using Remark 1 following Definition 5.3).

This proves Corollary 1.

COROLLARY 2. *Additivity of finiteness. If each of* $[0, a]$, $[0, b]$ *is finite, so is* $[0, a + b]$.

Proof. If $\{a_n\}$ is a homogeneous sequence in $[0, a + b]$, then $\{a_n\}$ is of type (A), and also of type (A^*) by Theorem 6.3 and its Corollary 1. Then, by Corollary 2 to Theorem 5.1, all a_n are 0.

Then, by the Corollary to Theorem 6.1, $[0, a + b]$ is finite.

7. Unrestricted additivity of continuity in finite lattices.

We assume that L is a complemented $\aleph$-complete modular lattice.

LEMMA 7.1. *Suppose* L *is upper* $\aleph$*-continuous. Then for every family* $\{a_\beta; \beta < \Omega\}$ *there exists a strongly independent family* $\{\bar{a}_\beta\}$ *such that* $\bar{a}_\beta \leqslant a_\beta$ *and* $\sum \bar{a}_\beta = \sum a_\beta$.

Proof. Put $\bar{a}_\beta = [a_\beta - a_\beta \sum (a_\gamma; \gamma < \beta)]$. Obviously $\{\bar{a}_\beta\}$ is independent, $\bar{a}_\beta \leqslant a_\beta$ and by transfinite induction on γ, $\sum (\bar{a}_\beta; \beta < \gamma) = \sum (a_\beta; \beta < \gamma)$ for all $\gamma < \Omega$.

Strong independence of $\{\bar{a}_\beta\}$ is equivalent to independence of $\{\bar{a}_\beta\}$ since L is upper $\aleph$-continuous (see the last sentence preceding Theorem 3.1).

COROLLARY. *If x is an $\aleph$-residual element and $[0, x]$ is upper $\aleph$-continuous, then x can be doubled in L.*

Proof. By the Corollary to Theorem 3.2 there exists a residual C-system $\{(x_\beta, c_\beta, b_\beta)\}$ with $x = \sum x_\beta$.

By Lemma 7.1, $\bar{x}_\beta = [x_\beta - x_\beta \sum (x_\gamma; \gamma < \beta)]$ has the properties: $\{\bar{x}_\beta\}$ is strongly independent, $\bar{x}_\beta \leqslant x_\beta$ and $x = \sum \bar{x}_\beta$.

Let $\bar{b}_\beta = b_\beta(\bar{x}_\beta + c_\beta)$. Then $\{\bar{b}_\beta\}$ is strongly independent (along with $\{b_\beta\}$), $\bar{x}_\beta \sim \bar{b}_\beta$ for each β and $(\sum \bar{x}_\beta)(\sum \bar{b}_\beta) \leqslant x \sum b_\beta = 0$. Now by (i) of the Corollary to Theorem 3.4, $x \sim \sum \bar{b}_\beta$. Since $x \sum \bar{b}_\beta = 0$ this proves that x can be doubled in L.

LEMMA 7.2. *Suppose $x \leqslant y$, $y \oplus Y = 1$ with $[0, Y]$ upper $\aleph$-continuous. If there exists an increasing family $\{c_\beta\}$ with $xc_\beta = 0$ for every β and $x \leqslant \sum c_\beta$, then there exists such an increasing family with $\sum c_\beta \leqslant y$.*

Proof. Let $c_\beta' = c_\beta + [1 - \sum c_\gamma]$. Then $\sum c_\beta' = 1$. Now yc_β' has the properties specified:

$$\sum (yc_\beta') = y \geqslant x$$

(observe that $[y, 1]$ is lattice isomorphic to $[0, Y]$ by (2.1), hence upper $\aleph$-continuous, and use Theorem 4.2);

$$x(yc_\beta') = xc_\beta' = x(\sum c_\gamma)c_\beta' = x(c_\beta + 0) = 0.$$

COROLLARY. *Suppose x is an $\aleph$-residual element in L and L is $\aleph'$-continuous for every $\aleph' < \aleph$. If $y \oplus Y = 1$ with $x \leqslant y$ and $[0, Y]$ upper $\aleph$-continuous then x is an $\aleph$-residual element in $[0, y]$.*

Proof. By hypothesis, x is the residual element of some residually independent family $\{a_\beta; \beta < \Omega\}$.

Define $c_\beta = \sum (a_\gamma; \gamma < \beta)$. Then $xc_\beta = 0$ for each β since L is upper $\aleph'$-continuous for all $\aleph' < \aleph$. And $x \leqslant \sum c_\beta$ since $\sum c_\beta = \sum a_\beta \geqslant x$.

Hence Lemma 7.2 shows that an increasing family $\{\bar{c}_\beta\}$ exists with $\sum \bar{c}_\beta \leqslant y$, $x\bar{c}_\beta = 0$ for every β and $x \leqslant \sum \bar{c}_\beta$. By Theorem 3.6, applied to $[0, y]$, $x \leqslant t$ for some t which is an $\aleph$-residual element in $[0, y]$; hence x itself has this property, by Theorem 3.3.

THEOREM 7.1. *Suppose x is an $\aleph$-residual element in L with $[0, x]$ upper $\aleph$-continuous. If L is $\aleph'$-continuous for all $\aleph' < \aleph$ then x is a member of a homogeneous sequence.*

Proof. 1. By Theorem 6.1 it is sufficient to show that x is a member of an independent sequence of pairwise perspective elements.

2. It is sufficient therefore to prove that if $\{x, x_1, \ldots, x_m\}$ is an independent family of pairwise perspective elements and $m \geqslant 0$, then there exists some $x_{m+1} \sim x$ such that $\{x, x_1, \ldots, x_{m+1}\}$ is independent.

3. We choose $Y = x_1 + \ldots + x_m$ and $y = [1 - Y]$ with $y \geqslant x$.

Since each $[0, x_i]$ is upper $\aleph$-continuous, along with $[0, x]$, so is $[0, Y]$, by Theorem 4.3. Hence, by the Corollary to Lemma 7.2, x is a residual element in $[0, y]$. Now the Corollary to Lemma 7.1, applied to $[0, y]$ shows that $x \sim x_{m+1}$ for some $x_{m+1} \leqslant y$ with $xx_{m+1} = 0$. Then

$$x_{m+1}(x + x_1 + \ldots + x_m) = x_{m+1}y(x + x_1 + \ldots + x_m) = x_{m+1}(x + 0) = 0$$

so $\{x, x_1, \ldots, x_{m+1}\}$ is independent.

This x_{m+1} satisfies our requirements and this completes the proof of Theorem 7.1.

COROLLARY 1. *If L is finite and locally $\aleph$-continuous then L is $\aleph$-continuous.*

Proof. We prove this by transfinite induction. Hence we can suppose L is $\aleph'$-continuous for all $\aleph' < \aleph$.

By Corollary 7 to Theorem 3.6 it is sufficient to show that every $\aleph$-residual element t must be 0.

But if $t \neq 0$ then for some non-zero x with $x \leqslant t$, $[0, x]$ is $\aleph$-continuous (*a fortiori*, upper $\aleph$-continuous). Then, by Theorem 3.3, x is also an $\aleph$-residual element. Now Theorem 7.1 shows that x is a member of a homogeneous sequence.

But L is finite, so $x = 0$. This gives a contradiction and shows that $t \neq 0$ is impossible. Thus Corollary 1 must be valid.

Remark. The proof of Theorem 7.1 shows that if L is finite and locally upper $\aleph$-continuous and $\aleph'$-continuous for all $\aleph' < \aleph$, then L is upper $\aleph$-continuous. When $\aleph = \aleph_0$ this becomes: if L is finite and locally upper $\aleph_0$-continuous then L is upper $\aleph_0$-continuous.

COROLLARY 2. *If in L every homogeneous sequence is of type* (A) *and L is locally upper $\aleph_0$-continuous then L is upper $\aleph_0$-continuous.*

(Note: the Remark following Corollary 1 to Theorem 7.1 uses the stronger condition that L be *finite*.)

Proof. Suppose, if possible, that L is not upper $\aleph_0$-continuous. Then there exists some $t \neq 0$ with t an $\aleph_0$-residual element in L.

By the hypotheses, there exists an $x \neq 0$ with $x \leqslant t$ and $[0, x]$ upper $\aleph_0$-continuous. By Theorem 3.3, x itself is also an $\aleph_0$-residual element in L.

Now Theorem 7.1 shows that x is a member of a homogeneous sequence $\{x, y_0, y_1, \ldots\}$, by the hypotheses necessarily of type (A).

Choose $Y = \sum y_n$ and $y = [1 - Y]$ with $y \geqslant x$. Then $[0, Y]$ is upper $\aleph_0$-continuous by Theorem 5.1 so x is an $\aleph_0$-residual element in $[0, y]$, by the Corollary to Lemma 7.2.

Now by the Corollary to Theorem 3.2 there exists a residual C-sequence $\{(x|c_n, b_n)\}$ with $x + \sum b_n \leqslant y$. Since b_n is perspective to a subelement of x and x is perspective to y_n and $(b_n + x)y_n \leqslant yY = 0$ therefore (2.2) shows that b_n is perspective to a subelement of y_n.

Since $x \sim y_0$ and $(x + \sum b_n)(\sum y_n) \leqslant yY = 0$, therefore (i) of the Corollary to Theorem 3.4 shows that $x + \sum b_n$ is perspective to a subelement of $\sum y_n$.

But $[0, \sum y_n] = [0, Y]$ is upper $\aleph_0$-continuous so $[0, x + \sum b_n]$ has the same property. Hence $x = x\sum c_n = \sum_{n=1}^{\infty}(x\sum_{i=1}^{n} c_i) = \sum(0) = 0$, a contradiction.

Thus Corollary 2 must be vaild.

8. Homogeneous sequences (continued). In this section we assume that L is a complemented countably complete modular lattice.

LEMMA 8.1. *Suppose* $\{a_n\}$ *is a homogeneous sequence with* $(a_n, x_n, a_{n+1})C$ *for every* n. *Then*

(i) $a_1(\sum x_n) = 0$ *implies* $\{a_n\}$ *is of type* (A);

(ii) $a_1 \leqslant \sum x_n$ *implies* $\{a_n\}$ *is of type* (A*).

Proof of (i): We shall show that $\sum x_i$ is a complement of every a_n in $\sum a_i$. Clearly $a_n + \sum x_i = \sum a_i$ so we need only prove $a_n\sum x_i = 0$ for each n. But the axis $x_1 + \ldots + x_{n-1}$ gives a perspective mapping of $[0, a_n]$ onto $[0, a_1]$ and by this mapping $a_n\sum x_i$ is mapped on $(a_n\sum x_i + x_1 + \ldots + x_{n-1})a_1 \leqslant (\sum x_i)a_1 = 0$, so $a_n\sum x_i$ itself must be 0.

Proof of (ii): We need only show that $\{x_n\}$ is a homogeneous sequence of type (A*). For the Corollary 1 to Theorem 5.1 will show that $[0, \sum x_n]$ is lower $\aleph_0$-continuous; then $[0, \sum a_n]$ will also be lower $\aleph_0$-continuous since the hypothesis implies that $\sum a_n \leqslant \sum x_n$; then $\{a_n\}$ will be of type (A*), again by Corollary 1 to Theorem 5.1.

To show that $\{x_n\}$ is homogeneous of type (A*) it is sufficient (by Remark 1 following Definition 5.3) to prove:

(8.1) $\{x_n\}$ is strongly independent,

(8.2) a_1 is a complement of every x^*_n in $\sum x_i$.

Now (8.1) follows from (3.1).

To prove (8.2) we verify:

$$a_1x_n^* = a_1\Big(\sum_{i=1}^{n} a_i\Big)\Big(\sum_{i=1}^{n-1} x_i + \sum_{i=n+1}^{\infty} x_i\Big)$$

$$= a_1 \sum_{i=1}^{n-1} x_i + 0 = 0;$$

$$a_1 + x_n^* = a_1 + x_1 + \ldots + x_{n-1} + \sum_{i=n+1}^{\infty} x_i$$

$$= a_1 + \ldots + a_{n-1} + \sum_{i=n+1}^{\infty} x_i$$

so we need only show

(8.3) $a_n \leqslant \sum_{i=n+1}^{\infty} x_i$ for every n.

But (8.3) holds for $n = 1$, by hypothesis. Also, (8.3) holds for $n = 2$ since

$$a_2 < \left(\sum_{i=2}^{\infty} a_i\right)(a_1 + x_1) < \left(\sum_{i=2}^{\infty} a_i\right)\left(\sum_{i=2}^{\infty} x_i + x_1\right)$$

$$= \left(\sum_{i=2}^{\infty} x_i\right) + \left(\sum_{i=2}^{\infty} a_i\right)x_1(a_1 + a_2)$$

$$= \left(\sum_{i=2}^{\infty} x_i\right) + 0 \qquad \text{since } x_1 a_2 = 0,$$

$$a_2 < \sum_{i=2}^{\infty} x_i.$$

By repetition of this calculation, (8.3) can be proved for all n. This shows that (ii) holds and completes the proof of Lemma 8.1.

THEOREM 8.1. (i) *If* $\{b_n\}$, $\{c_n\}$ *are homogeneous sequences of types* (A) *and* (A*) *respectively then* $\sum b_n$ *and* $\sum c_n$ *are completely disjoint.*

(ii) *If* $\{a_n\}$ *is a homogeneous sequence there is a unique decomposition* $a_n = b_n \oplus c_n$ *such that* $\{b_n\}$ *and* $\{c_n\}$ *are homogeneous sequences of types* (A), (A*) *respectively.*

Proof of (i): Let $b = \sum b_n$, $c = \sum c_n$. We may suppose $(b, c)P$ false and we need only derive a contradiction. Clearly we may suppose (replacing b_n, c_n by suitable subelements) $b_1 \sim c_1 \neq 0$. Then $x < c$, $(x, b)P$ together imply $x = 0$.

Let $d = [c - bc]$. Lemma 5.2 (with our d in place of A in Lemma 5.2) shows that $[0, d]$ is upper $\aleph_0$-continuous. And $[0, bc]$ is upper $\aleph_0$-continuous since $bc < b$ and $[0, b]$ is upper $\aleph_0$-continuous. Then Theorem 4.3 shows that $[0, c]$ is upper $\aleph_0$-continuous.

Hence, by Theorem 5.1, $\{c_n\}$ is of type (A). Since $\{c_n\}$ is also of type (A*), all c_n are 0 (by Corollary 2 to Theorem 5.1). This contradicts $c_1 \neq 0$ and (i) is therefore established.

Proof of (ii): We need only obtain one decomposition as described, since uniqueness will follow from (i).

Since $\{a_n\}$ is a homogeneous sequence, $(a_n, x_n, a_{n+1})C$ holds for certain x_n. Put $x = \sum x_n$ and let

$$c_1 = a_1 x, \qquad b_1 = [a_1 - a_1 x]$$

and for $n > 1$,

$$c_{n+1} = (c_n + x_n)a_{n+1} \qquad b_{n+1} = (b_n + x_n)a_{n+1}.$$

Thus, c_{n+1}, b_{n+1} obtain from c_n, b_n respectively by the perspective mapping of $[0, a_n]$ onto $[0, a_{n+1}]$ with axis x_n.

It follows that $\{b_n\}$, $\{c_n\}$ are homogeneous sequences and $a_n = b_n \oplus c_n$.

Now let

$$y_n = x_n(b_n + b_{n+1}), \qquad z_n = x_n(c_n + c_{n+1}).$$

Then as the reader can verify easily, $(b_n, y_n, b_{n+1})C$, $(c_n, z_n, c_{n+1})C$ hold and $x_n = y_n \oplus z_n$.

Since $b_1 \sum y_n \leqslant b_1 a_1 \sum x_n = 0$, Lemma 8.1 (i) shows that $\{b_n\}$ is a homogeneous sequence of type (A).

Since

$$\begin{aligned} c_1 = c_1 a_1 x &= c_1(\sum y_n + \sum z_n) \\ &= c_1(\sum y_n + (c_1 + \sum y_n)\sum z_n) = c_1 \sum y_n \end{aligned}$$

(observe: $(c_1 + \sum y_n)\sum z_n \leqslant (\sum c_n)(\sum b_n) = 0$), therefore Lemma 8.1 (ii) shows that $\{c_n\}$ is of type (A*).

This completes the proof of Theorem 8.1.

COROLLARY. *If L is complete then L is a direct sum $L_1 + L_2 + L_3$ where $L_i = [0, a_i]$ with all a_i in the centre of L, L_1 is upper $\aleph_0$-continuous, L_2 is lower $\aleph_0$-continuous and L_3 is finite (a_1, a_2, a_3 are unique if L_3 is maximal with the finiteness property).*

Proof. Let $a_1 = \sum x$ where x varies over all elements perspective to members of homogeneous sequences of type (A), let $a_2 = \sum y$ where y varies over all elements perspective to members of homogeneous sequences of type (A*). Then a_1, a_2 are in the centre of L.

Let $a_3 = [1 - (a_1 + a_2)]$. Then $[0, a_3]$ is clearly finite.

Now every homogeneous sequence in $[0, a_1]$ is of type (A). For otherwise $a_1 \neq 0$ and some $y \neq 0$ with $y \leqslant a_1$ would be a member of a homogeneous sequence of type (A*), by Theorem 8.1 (ii). Also $(y, x)P$ would be false for some x perspective to a member of a homogeneous sequence of type (A) hence for some x which is itself a member of a homogeneous sequence of type (A). But Theorem 8.1 (i) shows that $(y, x)P$ holds in such circumstances. This contradiction proves that every homogeneous sequence in $[0, a_1]$ is of type (A).

Also $[0, a_1]$ is locally upper $\aleph_0$-continuous. For if $y \neq 0$, $y \leqslant a_1$ holds, then as in the preceding paragraph $(y, x)P$ is false for some x which is a member of a homogeneous sequence of type (A), so $[0, x]$ is upper $\aleph_0$-continuous (using Theorem 5.1). Then for some $y_1 \neq 0$, $y_1 \leqslant y$, the lattice $[0, y_1]$ is perspective to $[0, x_1]$ for some $x_1 \leqslant x$ so $[0, y_1]$ is also upper $\aleph_0$-continuous. This proves $[0, a_1]$ to be locally upper $\aleph_0$-continuous.

Then Corollary 2 to Theorem 7.1 shows that $[0, a_1]$ is upper $\aleph_0$-continuous. Similarly, using the dual to Corollary 2 to Theorem 7.1, $[0, a_2]$ is lower $\aleph_0$-continuous.

Remark 1. If L is $\aleph$-complete but not necessarily complete we can show that L is a sublattice of such a direct sum $L_1 + L_2 + L_3$.

Remark 2. If L is complete and irreducible then L must be upper $\aleph_0$-continuous, or lower $\aleph_0$-continuous or finite.[13] (See the Note added at end of this paper.)

9. Kaplansky's theorem.

THEOREM 9.1. *Suppose L is a complemented countably complete modular lattice.*[14] *Then L is finite if it has the property:*

(9.1) *for every $a \neq 0$ there exists an anti-automorphism ϕ of L such that $(a, b)P$ is false for every complement b of $\phi(a)$.*[15]

Proof.[16] By the Corollary to Theorem 6.1 it is sufficient to show that if $\{a_n\}$ is a homogeneous sequence in L then $a_1 = 0$. By Theorem 8.1 (ii) we may suppose that $\{a_n\}$ is of type (A) or (A*).

Suppose if possible that $a_1 \neq 0$. Let ϕ be an anti-automorphism of L (as provided by (9.1)) such that $(a_1, b)P$ is false whenever $b \oplus \phi(a_1) = 1$. Then $\{\phi(a_n)\}$, considered in L', is homogeneous of type (A) or (A*).

Hence every dual sequence $\{b_n\}$ of $\{\phi(a_n)\}$ is homogeneous of type (A*) or (A) respectively, in L. Therefore $(a_1, b_1)P$ holds by Theorem 8.1 (i).

But b_1 is a complement of $\phi(a_1)$ by the definition of dual sequence, so this gives a contradiction to the property assumed for ϕ.

Thus $a_1 \neq 0$ is not possible and Theorem 9.1 is proved.

Remark 1. (9.1) is obviously implied by the property:

(9.2) *for every $a \neq 0$ there exists an anti-automorphism ϕ of L such that $a \leqslant \phi(a)$ is false.*

Hence, *if L is a countably complete, orthocomplemented modular lattice, then L must be finite* (see the Appendix for a direct proof of this result).

Remark 2. If L is a complemented, *complete* modular lattice, then (9.1) is implied by the property:

(9.2)′ *for every $z \neq 0$ with z in the centre of L, there exists an anti-automorphism ϕ of L such that $z \leqslant \phi(z)$ is false.*

To derive (9.1) *from:* (9.2)′ suppose $a \neq 0$ and let z be the least central element with $z \geqslant a$. Then there exists an anti-automorphism ϕ of L (as provided by (9.2)′) such that $z \leqslant \phi(z)$ is false.

[13] Irreducibility for a lattice L means: $L = L_1 + L_2$ (direct sum) only if L_1 or L_2 consists of one element. If L is complemented and modular, this is equivalent to: 0, 1 are the only elements in the centre of L (it was shown first by von Neumann (3, Part I, Theorems 5.2, 5.3) that for complemented modular lattices, irreducibility in the above sense is equivalent to: 0, 1 are the only elements with unique complements).

[14] Theorem 9.1 (and also its Corollary) may fail to hold if L is not countably complete. This failure occurs in the orthocomplemented modular lattice consisting of all the linear subspaces of finite dimension and their orthogonal complements in Hilbert space.

[15] If $(a, b)P$ is false for *one* complement b of $\phi(a)$ then $(a, b)P$ is necessarily false for every complement b of $\phi(a)$.

[16] See footnote 23.

Since $z \leqslant z_1$ for every central element $z_1 \geqslant a$ the anti-automorphism ϕ yields: $\phi(z) \geqslant z_1$ for every central element $z_1 \leqslant \phi(a)$. Since $\phi(z) \geqslant z$ is false, therefore $z \leqslant \phi(a)$ must be false, also. Hence $zb \neq 0$ for some complement b of $\phi(a)$.

Using (2.8) we have: $(a, b)P$ is false for at least one complement b of $\phi(a)$, hence for every complement b of $\phi(a)$.[17] Thus (9.1) has been derived from (9.2)'.

COROLLARY.[18] *A complete complemented modular lattice L is necessarily finite if it possesses an anti-automorphism ϕ which is an orthocomplementation on the centre (that is, $\phi(z) \oplus z = 1$ for every central element z), in particular if L is irreducible*[19] *and possesses at least one anti-automorphism.*

Definition 9.1. For a lattice L the property $(SI)_{\aleph}$ shall mean:

$(SI)_{\aleph}$: For every increasing family $\{c_\beta; \beta < \Omega\}$ there exists a strongly independent family $\{a_\beta; \beta < \Omega\}$ such that $c_\beta = \sum(a_\gamma; \gamma \leqslant \beta)$ for all $\beta < \Omega$.

THEOREM 9.2. *An orthocomplemented $\aleph$-complete modular lattice has the property $(SI)_{\aleph}$.*

Proof. Suppose $a \to \phi(a)$ denotes the orthocomplementation. If $\{c_\beta\}$ is an increasing family, choose $a_1 = c_1$, and for $1 < \beta < \Omega$, $a_\beta = c_\beta \prod(\phi(c_\gamma); \gamma < \beta)$.

Then $\{a_\beta\}$ is strongly independent; for $\gamma < \beta$ implies that a_γ is orthogonal to a_β; hence a_γ is orthogonal to a^*_γ, $\prod a^*_\gamma$ is orthogonal to every a_β, hence to $\sum a_\beta$. Since $\prod a^*_\gamma \leqslant \sum a_\beta$ this implies $\prod a^*_\gamma = 0$, so $\{a_\beta\}$ is indeed strongly independent.

By transfinite induction on β it is easy to show that $c_\beta = \sum(a_\gamma; \gamma \leqslant \beta)$ for all $\beta < \Omega$.

This proves Theorem 9.2.

THEOREM 9.3. *Suppose L is a complemented $\aleph$-complete modular lattice with the property $(SI)_{\aleph}$. If L is finite and can be doubled then L is upper $\aleph$-continuous.*

Proof 1. We may suppose $xc_\beta = 0$ for all $\beta < \Omega_1$ for some $\Omega_1 \leqslant \Omega$ and $c_\beta \uparrow 1$ and we need only prove $x = 0$.

Let X be a complement of x and let $c_\beta' = (x + c_\beta)X$. Then $[0, c_\beta]$ is mapped onto $[0, c_\beta']$ by the perspective mapping with axis x.

Since L is assumed to have the property $(SI)_{\aleph}$, there exist strongly independent families $\{a_\beta\}$, $\{a_\beta'\}$ such that $c_\beta = \sum(a_\gamma; \gamma \leqslant \beta)$, $c_\beta' = \sum(a_\gamma'; \gamma \leqslant \beta)$ for all $\beta < \Omega_1$.

Then $a_\beta \sim (x + a_\beta)X$ (with axis x). Since $a_\beta \oplus \sum(a_\gamma; \gamma < \beta) = c_\beta$, that is, $a_\beta \oplus \sum(c_\gamma; \gamma < \beta) = c_\beta$ therefore (by the perspective mapping with axis x),

[17] See footnote 15.
[18] See footnote 14.
[19] See footnote 13.

$$(x + a_\beta)X \oplus \sum(c_\gamma'; \gamma < \beta) = c_\beta'$$

so $(x + a_\beta)X \sim a_\beta'$ (with axis $\sum(a_\gamma'; \gamma < \beta) = \sum(c_\gamma'; \gamma < \beta)$).

Therefore $a_\beta \simeq a_\beta'$ for each $\beta < \Omega_1$, $\sum a_\beta = \sum c_\beta = 1 > \sum a_\beta'$ and L can be doubled.

If $x \neq 0$, then $\sum a_\beta' \leqslant X < 1$ so $[\sum a_\beta - \sum a_\beta'] \neq 0$ and by (ii) of the Corollary to Theorem 3.4, $[\sum a_\beta - \sum a_\beta']$ is a member of an independent sequence of pairwise perspective elements. This would contradict the assumed finiteness of L.

Hence $x = 0$ as required and Theorem 9.3 is proved.

THEOREM 9.4. *If L is orthocomplemented $\aleph$-complete and modular then L is $\aleph$-continuous.*[20]

Proof. L is finite by Remark 1 following Theorem 9.1. Thus, by Corollary 1 to Theorem 7.1 it is sufficient to prove local $\aleph$-continuity of L. Since L possesses an anti-automorphism it is sufficient to prove that L is locally upper $\aleph$-continuous.

We may suppose x is an element of L with $x \neq 0$ and we need only show that for some element y, with $0 \neq y \leqslant x$, the lattice $[0, y]$ is upper $\aleph$-continuous

Now if there exists an element $0 \neq y \leqslant x$ such that $[0, y]$ can be doubled then, since $[0, y]$ has the property $(\mathrm{SI})_\aleph$, Theorem 9.3 shows that $[0, y]$ is upper $\aleph$-continuous.

On the other hand, if $0 \neq y \leqslant x$ implies that $[0, y]$ cannot be doubled then $(y, z)P$ holds whenever $y \leqslant x$, $z \leqslant x$ with $yz = 0$.[21] Now if $\{a_\beta\}$ is an increasing family in $[0, x]$ and $ya_\beta = 0$ for every β, then $(y, a_\beta)P$ holds for every β, so $(y, \sum a_\beta)P$ holds, hence $y\sum a_\beta = 0$. This proves that $[0, x]$ is itself upper $\aleph$-continuous.

This completes the proof of Theorem 9.4.

THEOREM 9.5. *Suppose L is a complemented $\aleph$-complete modular finite lattice which possesses an anti-automorphism ϕ of period two with the following continuity property: $\phi(x_\beta) \oplus x_\beta = 1$, $x_\beta \uparrow x$ together imply $\phi(x) \oplus x = 1$. Then L is $\aleph$-continuous.*

Remark. Such ϕ generalize orthocomplementation.

Proof 1. We prove this theorem by transfinite induction on $\aleph$ so we may suppose L is $\aleph'$-continuous for all $\aleph' < \aleph$.

---

[20] This is a strengthened form of Kaplansky's theorem (1). In a letter to one of us dated June 13, 1957, Kaplansky conjectured that any complemented complete modular lattice is continuous if it possesses an anti-automorphism of period two which is an orthocomplementation on the centre. Our Theorem 9.1 establishes finiteness under even weaker conditions but our Theorem 9.5 establishes continuity only under conditions somewhat more restrictive than those of Kaplansky's conjecture.

[21] In this case $[0, x]$ is a complemented modular lattice in which every element has a unique complement, that is, a Boolean algebra.

2. It is sufficient to prove that L is locally upper $\aleph$-continuous, for the anti-automorphic character of L will show then that L is locally lower $\aleph$-continuous, hence locally $\aleph$-continuous. Since L is assumed to be finite, Corollary 1 to Theorem 7.1 will then show that L is $\aleph$-continuous.

3. Thus we may suppose $a \neq 0$ and we need only prove that $[0, b]$ is upper $\aleph$-continuous for some $0 \neq b < a$. We shall prove below:

(i) If $\phi(x) \geq x$ is false for every $0 \neq x < a$ then $[0, a]$ possesses an orthocomplementation $u \to a\phi(u)$ (Theorem 9.4 then shows that $[0, a]$ itself is $\aleph$-continuous).

(ii) If $\phi(x) \geq x$ holds for some $0 \neq x < a$ then $[0, x]$ is upper $\aleph$-continuous.

4. *To prove* (i): We note that

$$\phi(x\phi(x)) = \phi(x) + x \geq x\phi(x).$$

Since $x\phi(x) \leq x$ the assumption of (i) implies that $x\phi(x) = 0$ for all $x < a$. Then also $x + \phi(x) = 1$. Thus $x + a\phi(x) = a$, $a\phi(a\phi(x)) = a(\phi(a) + x) = x + a\phi(a) = x$ so $x \to a\phi(x)$ is an orthocomplementation on $[0, a]$. This proves (i).

5. *To prove* (ii): We may suppose $a_\beta \uparrow x (\beta < \Omega)$, $ya_\beta = 0$ for all β and $y \leq x$; we need only prove $y = 0$.

We can choose b_β by transfinite induction on β so that $b_1 = [x - a_1]$, and for $\beta > 1$,

$$b_\beta = [\prod(b_\gamma; \gamma < \beta) - a_\beta \prod(b_\gamma; \gamma < \beta)],$$

and so that $b_\beta \geq y$ for all β.[22] Then $a_\beta \oplus b_\beta = x$ for all β since $[0, x]$ is lower $\aleph'$-continuous for all $\aleph' < \aleph$.

Now let X be a complement of x. Set $c_\beta = X\phi(a_\beta)$. Then $c_\beta \downarrow$, $c_\beta x \leq Xx = 0$ and $c_\beta \oplus x = X\phi(a_\beta) + x = \phi(a_\beta)(X + x) = \phi(a_\beta)$ for all β. Hence

$$(b_\beta \oplus c_\beta) \oplus a_\beta = \phi(a_\beta), \qquad \phi(b_\beta + c_\beta)\phi(a_\beta) = a_\beta,$$
$$(b_\beta + c_\beta)\phi(b_\beta + c_\beta) = (b_\beta + c_\beta)\phi(a_\beta)\phi(b_\beta + c_\beta) = (b_\beta + c_\beta)a_\beta = 0,$$

so

$$(b_\beta + c_\beta)\phi(b_\beta + c_\beta) = 0$$

for all β.

Suppose $\phi(b_\beta + c_\beta) \uparrow d$. Then $d\phi(d) = 0$ and $(b_\beta + c_\beta) \downarrow \phi(d)$. Since $y \leq b_\beta$ for every β, therefore $y \leq \prod(b_\beta + c_\beta)$, that is, $y \leq \phi(d)$; but also

$$\phi(d) \leq \prod(b_\beta + c_\beta) \leq \prod\phi(a_\beta) = \phi(\sum a_\beta) = \phi(x)$$

so $d \geq x$. Thus $y \leq x\phi(d) \leq d\phi(d) = 0$ as required.

This proves that $[0, x]$ is upper $\aleph$-continuous and completes the proof of Theorem 9.5.

---

[22] Choose
$$b_1 = y + [x - (y + a_1)],$$
$$b_\beta = y + [\prod(b_\gamma; \gamma < \beta) - (y + a_\beta) \prod(b_\gamma; \gamma < \beta)].$$

Appendix on Finiteness

THEOREM.[23] *In an orthocomplemented countably complete modular lattice every independent*[24] *sequence of pairwise perspective elements must have all its elements zero.*

Proof. 1. We call an infinite sequence $\{x_n; n \geqslant 1\}$ *residually independent* if $x_n \sum (x_i; i \neq n) = 0$ for all n, *strongly independent* if $\prod_n \sum (x_i; i \neq n) = 0$. Since, for every p,

$$x_p \sum (x_i; i \neq p) \leqslant \prod_n \sum (x_i; i \neq n)$$

strong independence implies residual independence.

We note that if the x_n are pairwise orthogonal then $\{x_n\}$ is strongly independent. Also, if $\{x_n\}$ is strongly independent and for every n, $y_n \leqslant \sum (x_i; i \geqslant n)$, $y_n \sum (x_i; i > n) = 0$ then $\{y_n\}$ is also strongly independent. If $\{x_n\}$ is strongly independent then so is $\{(x^*_n)^\perp \sum x_n\}$ where x^*_n denotes $\sum (x_i; i \neq n)$ and $x^\perp$ denotes the orthogonal complement of the element x.

2. We may suppose $\{a_n\}$ is an independent sequence of pairwise perspective elements with $a_1 \neq 0$ and we need only derive a contradiction. By replacing a_n for $n > 1$ by $(a_1 + \ldots + a_n)(a_1 + \ldots + a_{n-1})^\perp$ we may even suppose that $\{a_n\}$ is strongly independent.

3. By using suitable replacements for the a_n we may even assume that they have a common relative complement A, that is, $a_n + A = \sum a_n$, $a_n A = 0$ for all n.[25] To see this, suppose $a_n \sim a_{n+1}$ (axis x_n), that is, $a_n + x_n = a_{n+1} + x_n = a_n + a_{n+1}$, $a_n x_n = a_{n+1} x_n = 0$. Let $x = \sum x_n$. We must consider two cases: $x a_1 \neq a_1$ and $x a_1 = a_1$.

---

[23] This theorem (first proved by Kaplansky (1, Theorem 1), see footnote 25) is contained in our Theorem 9.1 (see Remark 1 following Theorem 9.1) but we give here a direct (lattice-theoretic) proof for this orthocomplemented case which can be read independently of the rest of this paper provided the reader has some slight familiarity with complemented modular lattices.

With slight modification this direct proof actually establishes Theorem 9.1 in full generality.

[24] A family $\{x_\lambda; \lambda \in \Lambda\}$ is called independent if for every *finite* subset $F \subset \Lambda$, $x_\mu \Sigma (x_\lambda; \lambda \in F) = 0$ whenever $\mu \notin F$.

[25] Kaplansky constructs a common relative complement A for *every* sequence $\{a_n\}$ of pairwise orthogonal and perspective elements (of course, the Theorem will show finally that all a_n must be 0).

Kaplansky's method is as follows: first, he replaces L by $[0, \Sigma a_n[$. Then he shows that $\{\Sigma (a_{4n+i}; n \geqslant 1); i = 0, 1, 2, 3\}$ is a homogeneous basis of order 4 in the sense of von Neumann (3, Part II, Definition 3.1). Therefore L can be identified with the lattice of principal right ideals of a suitable regular ring $\mathfrak{R}$, by the coordinatization theorem of von Neumann (3, Part II, Theorem 14.1).

Since L is orthocomplemented, there exists a conjugation operation in $\mathfrak{R}$: $x \to x^*$ (that is, $(x + y)^* = x^* + y^*$, $(xy)^* = y^* x^*$ and $x^{**} = x$) such that every lattice element a in L is of the form $(e)_r$ with e a unique idempotent which is Hermitian (that is, $e^* = e$) and then $a^\perp = (1\text{-}e)_r$ (all proved in von Neumann (3, chapter II, Theorem 4.5)).

If $xa_1 \neq a_1$, let a_1' be a relative complement of xa_1 in a_1, let a_{n+1}' be obtained from a_n' by the perspective mapping with axis x_n, that is, $a_{n+1}' = (a_n' + x_n)a_{n+1}$ and let $x_n' = x_n(a_n' + a_{n+1}')$, $x' = \sum x_n'$. Then $\{a_n'\}$ is strongly independent, pairwise perspective and the a_n' have x as a common relative complement.

If $xa_1 = a_1$, then $\{x_n\}$ is strongly independent, pairwise perspective and the $x^*_n = \sum(x_i; i \neq n)$ have a_1 as common relative complement. Hence $\{(x^*_n)^\perp\}$ is strongly independent, pairwise perspective and $a_1^\perp$ is a common relative complement.

4. Suppose now that our strongly independent sequence of pairwise perspective elements is written as two sequences a_n, b_n, and let A be their common relative complement. The strong independence of the set {all a_n, all b_n} implies $(\sum a_n)(\sum b_n) = 0$.

Then $a_n \sim a_{n+1}$ with axis $c_n = A(a_n + a_{n+1})$ and $a_1 \sum c_n = 0$.

5. We shall show below that there exists a *residually* independent sequence of pairwise perspective elements $\{b_n'\}$ with $b_1' \sim b_1$, $(\sum b_n')(\sum a_n) = 0$ and axes of perspectivity d_n between b_n' and b_{n+1}' such that $b_1 \leqslant \sum d_n$.

6. We will then derive a contradiction as follows: $a_1 \sim b_1$, $b_1 \sim b_1'$ with $a_1(b_1 + b_1') = 0$ implies $a_1 \sim b_1'$. Let t_1 be an axis of perspectivity for a_1, b_1' and define t_{n+1} for $n \geqslant 1$, by induction as follows:

$$t_{n+1} = (t_n + c_n + d_n)(a_n + b_n').$$

Let $t = \sum t_n$. Then $t + \sum d_n = t + \sum c_n$ and (use: for each n, $\{a_n, \sum(a_m; m > n), b_n, \sum(b_m; m > n)\}$ is independent since $\{a_m\}$ and $\{b_m\}$ are both residually independent):

$$t \sum a_n = \left(t_1 + \sum_{i=2}^{\infty} t_i\right)\left(a_1 + \sum_{i=2}^{\infty} a_i\right) = t_1 a_1 + \left(\sum_{i=2}^{\infty} t_i\right)\left(\sum_{i=2}^{\infty} a_i\right)$$

$$= 0 + \left(\sum_{i=2}^{\infty} t_i\right)\left(\sum_{i=2}^{\infty} a_i\right) = \ldots \leqslant \Pi_n\left(\sum_{i=2}^{\infty} a_i\right) = 0.$$

Thus the given a_n must be of the form $(e_n)_r$ with all e_n idempotent, Hermitian and $e_n e_m = 0$ for $n \neq m$.

Kaplansky now constructs elements in $\mathfrak{R}$, namely w, e_{1i}, e_{i1} (for $i \geqslant 1$) such that: $e_{11} = e_1$; for all i, $e_{1i} = e_1 e_{1i} e_i$, $e_{1i} e_{i1} = e_1$ and $e_{i1} e_{1i} = e_i$; $w = e_1 w$ and $we_i = e_{1i}$ for all i (see Kaplansky (1, Lemma 21)).

Now $(w)^r$, the set of all u such that $wu = 0$, is a principal right ideal, as shown by von Neumann (3, Part II, Lemma 2.2).

This $(w)^r$ is a common complement of the a_n. For $w(e_{i1} - e_1) = 0$ for every i; so for every j, $(e_j)_r + (w)^r$ contains $e_j e_{j1} + (e_{j1} - e_1)(-e_1) = e_1$, hence it contains also, for every i, $(e_{i1} - e_1) + e_1 = e_{i1}$, hence also $e_{i1} e_{1i} = e_i$. Thus

$$(e_j)_r + (w)^r = 1 \text{ for all } j.$$

Finally, if u is in both $(e_j)_r$ and $(w)^r$ then $u = e_j u$ and $wu = 0$; that is, $we_j u = e_{1j} u = 0$, so $u = e_{j1}(e_{1j} u) = 0$. This means the meet of $(e_j)_r$ and $(w)^r$ is 0, and proves that $(w)^r$ is a common complement of all $a_n = (e_n)_r$.

Then

$$b_1' = b_1' \sum d_n = b_1'(t + \sum c_n)(t + a_1) = b_1'(t + (\sum c_n)a_1)$$
$$= b_1'(t + 0) = b_1' t = b_1' t_1 = 0.$$

This implies $a_1 = 0$ and gives the desired contradiction.

7. Thus we need only to construct $\{b_n'\}$ as described in 5 above. Since $A \sum b_m$ is a common relative complement of the b_n we may suppose $\sum b_m = 1$ and write A again in place of $A \sum b_m$. Now let a denote $A^\perp$ and let $\bar{b}_n$ denote $(b^*_n)^\perp$.

We shall prove that the family $\{a, \bar{b}_1, \bar{b}_2, \ldots\}$ are pairwise perspective and independent, $\bar{b}_1 \leqslant a + \sum(\bar{b}_i; i > 1)$, and $\bar{b}_1 \sim b_1$.

Indeed:

(i) $$(a\bar{b}^*_n)^\perp = A + \left(\sum_{m \neq n} \bar{b}_m\right)^\perp = A + \prod_{m \neq n} b^*_m = A + b_n = 1$$

since $\{b_n\}$ strongly independent implies that $\prod_{m \neq n} b^*_m = b_n$, and so $a\bar{b}^*_n = 0$.

(ii) $$(a + \bar{b}^*_n)^\perp = A b_n = 0, \text{ so } a + \bar{b}^*_n = 1.$$

Since $\{\bar{b}_n\}$ is strongly independent, (i) shows that $\{a, \bar{b}_1, \bar{b}_2, \ldots\}$ is independent. Then (i) and (ii) show that $a \sim \bar{b}_n$ with axis $\bar{b}^*_n$ so all of $\{a, \bar{b}_1, \bar{b}_2, \ldots\}$ are pairwise perspective, and $\bar{b}_1 \leqslant a + \bar{b}^*_1 = 1$. Finally, $\bar{b}_1 \sim b_1$ with axis b^*_1.

Thus the b_n' will be available, as described in 5. if we prove the following "orthogonalization" lemma.

LEMMA.[26] *Suppose $\{b_0, f_1, f_2, \ldots\}$ is independent and pairwise perspective. Then there exists a sequence $b_0, b_1, b_2, \ldots$ such that $\{b_n; n \geqslant 0\}$ is residually independent, and for $n \geqslant 0$, $b_{n-1} \sim b_n$ with axis d_n, so that $f_1 + \ldots + f_m = d_1 + \ldots + d_m$ for every m (in particular if $b_0 \leqslant \sum(f_m; m \geqslant 1)$ then $b_0 \leqslant \sum(d_m; m \geqslant 1)$).*

Proof. 1. Choose $d_1 = f_1$. Then b_0 and f_1 have some axis of perspectivity u and we choose $b_1 = (b_0 + f_1)u$. Then [27]

$$b_0 \oplus d_1 = b_1 \oplus d_1 = b_0 \oplus b_1.$$

Hence we can choose B_0, a complement of b_0, so that

$$B_0 = b_1 + [1 - (b_1 + d_1)]. \qquad \text{Let } B_{-1} = 1.$$

---

[26] This "orthogonalization" lemma is proved here for every complemented countably complete modular lattice. Even countable completeness need not be assumed if residual independence of $\{x_n\}$ is defined to mean: for every n there exists an element X_n such that $x_n X_n = 0$ and $X_n \geqslant x_m$ for all $n < m$.

[27] $x \oplus y$ denotes the lattice union $x + y$ but implies that the meet xy is 0. When $x \leqslant y$ the symbol $[x - y]$ denotes an arbitrary but fixed relative complement of y in x, that is, $[x - y] \oplus y = x$.

2. We may suppose that for some $r \geqslant 1$, the following statements hold:

$$(W)_r \begin{cases} d_1, \ldots, d_r, b_1, \ldots, b_r, B_0, \ldots, B_{r-1} \\ \text{have all been defined so that:} \\ b_{n-1} \sim b_n \text{ with axis } d_n \text{ for } n = 1, \ldots, r; \\ d_1 + \ldots + d_s = f_1 + \ldots + f_s \text{ for } s = 1, \ldots, r; \\ 1 = b_0 \oplus B_0; B_0 = b_1 \oplus B_1; \ldots; B_{r-2} = b_{r-1} \oplus B_{r-1}; \\ B_{r-1} \geqslant b_r, \end{cases}$$

and we need only show how to define d_{r+1}, b_{r+1}, B_r so that $(W)_{r+1}$ holds. (Observe that $(W)_1$ does hold for the d_1, b_1, B_0 defined in 1 above.)

3. Choose $d_{r+1} = B_{r-1}(f_1 + \ldots + f_{r+1})$.

Then

$$\begin{aligned} \text{(i)}\quad d_1 + \ldots + d_{r+1} &= (d_1 + \ldots + d_r + B_{r-1})(f_1 + \ldots + f_{r+1}) \\ &= (d_1 + \ldots + d_r + b_r + B_{r-1})(f_1 + \ldots + f_{r+1}) \\ &= (d_1 + \ldots d_r + b_{r-1} + B_{r-1})(f_1 + \ldots + f_{r+1}) \\ &= (d_1 + \ldots d_{r-1} + B_{r-2})(f_1 + \ldots + f_{r+1}) \\ &= \ldots = B_{-1}(f_1 + \ldots + f_{r+1}) = f_1 + \ldots + f_{r+1}. \end{aligned}$$

$$\begin{aligned} \text{(ii)}\quad & d_{r+1}(b_0 + b_1 + \ldots + b_r) = d_{r+1}b_r \\ & = d_{r+1}(f_1 + \ldots + f_{r+1})b_r(b_0 + f_1 + \ldots + f_r) \\ & = d_{r+1}b_r(f_1 + \ldots + f_r) = d_{r+1}b_r(d_1 + \ldots + d_r) = 0, \\ & d_{r+1}(f_1 + \ldots + f_r) \leqslant d_{r+1}(b_0 + b_1 + \ldots + b_r) = 0, \end{aligned}$$

so $d_{r+1} \sim f_{r+1}$ (axis $f_1 + \ldots + f_r = d_1 + \ldots + d_r$).

But $f_{r+1} \sim b_0$ and $b_0(f_{r+1} + d_{r+1}) \leqslant b_0(f_1 + \ldots + f_{r+1}) = 0$ so $d_{r+1} \sim b_0$.

Now $\{b_0, b_1, \ldots, b_r\}$ is independent, $b_{n-1} \sim b_n$ for $n \leqslant r$, and $d_{r+1}(b_0 + \ldots + b_r) = 0$, so $d_{r+1} \sim b_r$.

4. Since $d_{r+1} \sim b_r$, there exists an axis u such that

$$b_r \oplus u = d_{r+1} \oplus u.$$

We choose $b_{r+1} = (b_r \oplus d_{r+1})u$. Then we have

$$b_r \oplus d_{r+1} = b_{r-1} \oplus d_{r+1} = b_r \oplus b_{r+1}$$

so $b_r \sim b_{r+1}$ with axis d_{r+1}.

5. Since $b_r + d_{r+1} \leqslant B_{r-1}$ we can choose

$$B_r = b_{r+1} + [B_{r-1} - (b_r + d_{r+1})].$$

Then $B_r \geqslant b_{r+1}$, and $B_{r-1} = b_r \oplus B_r$.

Thus $(W)_{r+1}$ is satisfied and so the Lemma is proved and hence the Theorem is proved.

Note added in proof. A recent paper by Ornstein (*Dual vector spaces*, Ann. Math., *69* (1959), 520–34) obtains the following result (his Corollary 5.1):

Suppose L is a complete, atomic, centreless, complemented, modular lattice in which 1 is the union of a countable number of atoms and O is the intersection of a countable number of co-atoms; then L is either isomorphic or anti-isomorphic to the lattice of all subspaces of a vector space of countable dimension.

Ornstein's result can be deduced also from Remark 2 at the end of our § 8.

References

1. Irving Kaplansky, *Any orthocomplemented complete modular lattice is a continuous geometry*, Ann. Math., **61** (1955), 524–41.
2. F. Maeda, *Kontinuierliche Geometrien* (translation from the Japanese) (Berlin, Springer-Verlag, 1958).
3. J. von Neumann, *Lectures on continuous geometry*, Parts I, II, III, planographed (Princeton, The Institute For Advanced Study, 1935–7), to be reproduced in book form by Princeton University Press in 1959.
4. U. Sasaki, *On an axiom of continuous geometry*, J. Sci. Hiroshima Univ., Ser. A., **14** (1950), 100–1.

Queen's University
Tokyo College of Science

SUR UN THEORÈME DE G. SZEKERES, CONCERNANT LES FONCTIONS MONOTONES ET CONVEXES

S. MARCUS

1. Dans une note récente **(1)**, Szekeres a donné le théorème suivant:

Si f est une fonction réelle, définie, strictement croissante et ayant la deuxième dérivée continue sur (a, b) $(-\infty \leqslant a < b \leqslant \infty)$, alors il existe une fonction réelle ϕ, définie et strictement croissante sur (a, b), et une fonction réelle ψ, définie et strictement croissante sur $(\phi(a), \phi(b))$, telles que $\phi''(x) \geqslant 0$ pour $a < x < b$, $\psi''(u) \leqslant 0$ pour $\phi(a) < u < \phi(b)$ et $f(x) = \psi(\phi(x))$ pour $a < x < b$.

Dans la démonstration de ce théorème, Szekeres utilise une expression où la dérivée $f'(x)$ apparaît au dénominateur, ce qui exige qu'on ait $f'(x) \neq 0$ pour $a < x < b$, condition qui ne figure pas dans l'énoncé du théorème. Il est naturel alors de poser le problème si le théorème ci-dessus reste en vigueur si cette condition n'est pas remplie. Il y a lieu aussi de se demander si les autres conditions, par exemple l'existence et la continuité de $f''(x)$ sont nécessaires pour qu'une fonction strictement croissante puisse être représentée comme une superposition de deux fonctions croissantes, l'une convexe, l'autre concave. Dans ce qui suit nous allons donner une réponse à ces questions. Nous allons voir non seulement que le théorème cesse d'être vrai si $f'(x)$ s'annule (même en un seul point de (a, b)), mais qu'on a aussi d'autres conditions nécessaires, de nature assez restrictive, pour que f admette une représentation du type que nous venons de parler. D'autre part, nous allons établir des conditions suffisantes assez voisines de celles nécessaires, de sorte qu'une amélioration de ces conditions ne paraît pas possible sans compliquer les hypothèses—hors mis le cas où les conditions nécessaires établies par le théorème 1 sont aussi suffisantes, question que nous ne savons pas résoudre.

2. Une fonction réelle g est *convexe* (*concave*) sur (a, b) si pour $a < x < y < b$ on a

$$g\left(\frac{x+y}{2}\right) < \frac{g(x)+g(y)}{2} \quad \left(g\left(\frac{x+y}{2}\right) > \frac{g(x)+g(y)}{2}\right).$$

Considérons une fonction réelle f, définie sur (a, b) $(-\infty \leqslant a < b \leqslant \infty)$. Supposons qu'il existe une fonction réelle ϕ, définie, convexe et strictement croissante sur (a, b) et une fonction réelle ψ, définie, concave et strictement croissante sur $(\phi(a), \phi(b))$, telles que

$$f(x) = \psi(\phi(x)) \text{ pour } a < x < b.$$

Nous dirons en ce cas que f est une *fonction du type K*.

Reçu le 25 novembre, 1958.

Si pour une fonction réelle f il existe, en un point x, la dérivée à droite (gauche) finie, nous désignons cette dérivée par $f^+(x)(f^-(x))$.

THÉORÈME 1. *Si une fonction réelle f, définie sur (a, b), est du type K sur (a, b), alors les fonctions $f^+(x)$ et $f^-(x)$ sont définies sur (a, b).* Chacune de ces deux fonctions est dérivable presque partout sur (a, b) et la dérivée est sommable sur tout intervalle compact contenu dans (a, b). Pour $a < x < b$ on a $f^-(x) > 0 < f^+(x)$.

Démonstration. Les fonctions ϕ et ψ, étant monotones, admettent des points de continuité donc, en vertu de la convexité de ϕ et de la concavité de ψ, ϕ et ψ sont continues en chaque point de (a, b) resp. $(\phi(a), \phi(b))$. D'après un théorème classique, $\phi^-(x)$ et $\phi^+(x)$ existent en chaque point de (a, b), $\psi^-(u)$ et $\psi^+(u)$ existent en chaque point de $(\phi(a), \phi(b))$. On a donc, en vertu de la monotonie stricte de ϕ,

$$f^-(x) = \psi^-(\phi(x)) \cdot \phi^-(x), f^+(x) = \psi^+(\phi(x)) \cdot \phi^+(x) \text{ si } a < x < b.$$

En vertu de la convexité de ϕ, ϕ^-, et ϕ^+ sont des fonctions croissantes sur (a, b). En vertu de la concavité de ψ, ψ^- et ψ^+ sont des fonctions décroissantes sur $(\phi(a), \phi(b))$, donc $\psi^-(\phi(x))$ et $\psi^+(\phi(x))$ sont décroissantes sur (a, b). Par suite, chacune des fonctions f^- et f^+ est un produit de deux fonctions monotones, donc de deux fonctions dérivables presque partout. Il s'ensuit que f^- et f^+ sont dérivables presque partout sur (a, b).

Soit maintenant $\langle \alpha, \beta \rangle$ un intervalle compact contenu dans (a, b). $\psi^-(x)$, $\psi^-(\phi(x))$, $\phi^+(x)$ et $\psi^+(\phi(x))$ sont monotones et bornées sur $\langle \alpha, \beta \rangle$, donc f^- et f^+ sont à variation bornée sur $\langle \alpha, \beta \rangle$. Mais d'après un théorème bien connu, telles fonctions admettent des dérivées sommables sur $\langle \alpha, \beta \rangle$.

f, comme superposition de fonctions strictement croissantes, est strictement croissante sur (a, b) donc, en vertu des théorèmes connus (par exemple celui de Scheefer **(2**, p. 138**)**) f^+ ne peut s'annuler identiquement en aucun intervalle. Il s'ensuit que $\phi^+(x)$ et $\psi^+(\phi(x))$ ne peuvent s'annuler identiquement en aucun intervalle contenu dans (a, b) donc ϕ^+, qui d'après le raisonnement ci-dessus est croissante sur (a, b), est même *strictement* croissante sur (a, b). On constate aussi que $\psi^+(\phi(x))$ est *strictement* décroissante. D'autre part, en vertu de la monotonie croissante des fonctions ϕ et ψ, on a $\phi^+(x) > 0 < \psi^+(\phi(x))$ pour $a < x < b$ et, en vertu de la monotonie stricte de $\phi^+(x)$ et $\psi^+(\phi(x))$, que nous venons d'établir, on conclut: $\phi^+(x) > 0 < \psi^+(\phi(x))$ pour $a < x < b$.

D'une manière analogue on prouve que $\phi^-(x) > 0 < \psi^-(\phi(x))$ pour $a < x < b$ et le théorème 1 est complètement démontré.

Un exemple d'une fonction f qui satisfait toutes les hypothèses du théorème de Szekeres mais ne remplit pas la condition $f^-(x) \neq 0 \neq f^+(x)$ est donné par $f(x) = x^3$, considérée sur n'importe quel intervalle ouvert contenant l'origine. En effet, on a $f'(0) = 0$. Donc, en vertu du théorème 1, $f(x) = x^3$ n'est du

type K sur aucun intervalle (a, b) tel que $a < 0 < b$, c'est-à-dire cette fonction ne vérifie pas la thèse du théorème de Szekeres.

3. Maintenant se pose le problème: les conditions nécessaires données par le théorème 1 sont-elles aussi suffisantes pour que la fonction f soit du type K? Nous ne savons pas. Mais nous allons montrer qu'il suffit d'ajouter, aux conditions nécessaires établies par le théorème 1, la condition que f^+ soit dérivable non seulement presque partout, mais partout sur (a, b), pour que f soit du type K.

LEMME. *Si g est une fonction réelle croissante et absolument continue sur (a, b) et si ϕ est une fonction réelle absolument continue sur $(g(a), g(b))$, alors la fonction λ donnée par $\lambda(x) = \phi(g(x))$ est absolument continue sur (a, b).*

Nous laissons au lecteur la démonstration de ce lemme.

Rappelons qu'une fonction f est *absolument continue au sens généralisé* sur (a, b), si elle est continue sur (a, b) et s'il existe une décomposition $(a, b) = \infty' E_n$, où E_n sont des ensembles tels que f soit absolument continue sur E_n, quel que soit n (3, p. 223).

THÉORÈME 2. *Soit f une fonction réelle dénbie sur (a, b) $(-\infty \leqslant a < b \leqslant \infty)$. Supposons que f est dérivable sur (a, b) et que $f'(x) > 0$ pour $a < x < b$. Supposons encore que $f''(x)$ existe, finie, en chaque point de (a, b) et que f'' est sommable sur chaque intervalle compact contenu dans (a, b). Alors, f est, sur (a, b), une fonction du type K.*

Remarque. Pour comprendre la situation du théorème 2 par rapport au théorème 1 il faut rappeler le théorème de Dini, d'après lequel, si f^+ est continue en x, alors $f'(x)$ existe et est finie (voir, par exemple, (3, p. 204)).

Démonstration du théorème 2. On va utiliser l'idée de la démonstration du théorème 1 de (1).

La fonction

$$f''_+(x) = \begin{cases} f''(x) \text{ si } f''(x) > 0 \\ 0 \text{ si } f''(x) \leqslant 0 \end{cases}$$

est sommable sur chaque intervalle compact contenu dans (a, b). Par hypothèse, $f'(x)$ est continue et strictement positive sur (a, b). Il existe donc pour chaque intervalle compact $J \subset (a, b)$ un nombre $\rho(J) > 0$ tel que

$$\frac{1}{\rho(J)} < f'(x) < \rho(J)$$

pour $x \in J$. Il s'ensuit donc que l'intégrale

$$p(y) = \int_d^y \frac{f''_+(x)}{f'(x)}\, dx$$

où d est un point fixe de (a, b), existe, au sens de Lebesgue, pour chaque y tel que $a < y < b$.

La fonction donnée par $p(y)$ est continue, strictement croissante et strictement positive sur (a, b). Il s'ensuit donc que la fonction

$$\phi(x) = \int_a^x e^{p(y)}dy$$

où l'intégrale est prise au sens de Cauchy, est strictement croissante et convexe sur (a, b).

Désignons par ψ l'inverse de la fonction ϕ. ψ est une fonction définie sur $(\phi(a), \phi(b))$, strictement croissante et concave sur $(\phi(a), \phi(b))$ et ayant les valeurs dans (a, b). Le théorème 2 sera démontré dès qu'on aura prouvé que la fonction donnée par

$$h(u) = f(\psi(u))$$

est concave sur $(\phi(a), \phi(b))$, car on a en tout cas, en vertu de ce que ψ et ϕ sont des fonctions inverses,

$$h(\phi(x)) = f(\psi(\phi(x))) = f(x).$$

Mais ϕ est une fonction dérivable, à dérivée strictement positive sur (a, b), donc ψ est dérivable sur $(\phi(a), \phi(b))$. Il s'ensuit que h est une fonction dérivable sur $(\phi(a), \phi(b))$. Pour montrer que h est concave, il suffit de montrer que $h'(u)$ est décroissante sur $(\phi(a), \phi(b))$. Le théorème sera donc démontré si nous pourrons montrer que

1° $h'(u)$ est absolument continue sur tout intervalle compact contenu dans $(\phi(a), \phi(b))$;

2° $h''(u) \leqslant 0$ presque partout sur $(\phi(a), \phi(b))$.

En effet, dans ce cas $h'(u)$ sera absolument continue au sens généralisé sur $(\phi(a), \phi(b))$. Mais on sait d'après un théorème classique (voir, par exemple, (3, p. 223)) qu'une fonction absolument continue au sens généralisé est décroissante dès que sa dérivée approximative est presque partout $\leqslant 0$.

Démontrons donc, 1° et 2°. Remarquons d'abord que la dérivée de ϕ est une fonction absolument continue sur tout intervalle compact J contenu dans (a, b). En effet, on a pour chaque $x \in (a, b)$, $\phi'(x) = e^{p(x)}$, où $p(x)$ est croissante et, étant une intégrale de Lebesgue, est absolument continue sur J.

Remarquons ensuite que la borne inférieure de $\phi'(x)$ sur J est strictement positive. D'autre part, ψ étant continue et concave, est absolument continue sur $\phi(J)$. En tenant compte que ψ est croissante, on déduit, en vertu du lemme, que $\phi'(\psi(u))$ est absolument continue et strictement positive sur $\phi(J)$. Il s'ensuit donc que

$$\psi'(u) = \frac{1}{\phi'(\psi(u))}$$

est aussi absolument continue sur $\phi(J)$, qui peut être considéré comme un intervalle compact quelconque contenu dans $(\phi(a), \phi(b))$. $f'(x)$, ayant une dérivée finie sur J, est absolument continue au sens généralisé sur J. La fonction

$$F(x) = \int_a^x f''(u)du$$

où α est l'extrémité gauche de J, est absolument continue sur J. Mais les dérivées des fonctions $f'(x)$ et $F(x)$ coïncident presque partout sur J donc, en vertu d'un théorème connu (3, p. 225) ces fonctions diffèrent par une constante. Il s'ensuit que $f'(x)$ est absolument continue sur J.

ψ, étant continue et concave, est absolument continue sur tout intervalle compact contenu dans $(\phi(a), \phi(b))$. En vertu de ce que ψ est croissante, on peut faire usage du lemme et conclure que $f'(\psi(u))$ est absolument continue sur $\phi(J)$. De ce que $h'(u) = f'(\psi(u)) \cdot \psi'(u)$ on déduit alors que h' est absolument continue sur $\phi(J)$.

Démontrons maintenant que $h''(u) \leqslant 0$ presque partout sur $(\phi(a), \phi(b))$. L'existence, presque partout, de $h''(u)$, est assurée par le fait que $h'(u)$, étant absolument continue, est à variation bornée sur tout intervalle compact contenu dans $(\phi(a), \phi(b))$. On doit montrer que

$$h''(u) = f''(\psi(u)) \cdot \psi'^2(u) + f'(\psi(u)) \cdot \psi''(u) \leqslant 0$$

presque partout sur $(\phi(a), \phi(b))$. Mais on a

$$\psi''(u) = -\frac{\phi''(\psi(u))}{\phi'^3(\psi(u))}.$$

En posant $x = \psi(u)$ et en remarquant que chacune des fonctions ϕ et ψ, en vertu de leur continuité absolue, transforme chaque ensemble de mesure nulle dans un ensemble de mesure nulle (3, p. 225) l'inégalité ci-dessus devient

$$\frac{f''(x)}{\phi'^2(x)} - f'(x)\frac{\phi''(x)}{\phi'^3(x)} \leqslant 0$$

presque partout sur (a, b), donc

$$\frac{f''(x)}{f'(x)} \leqslant \frac{\phi''(x)}{\phi'(x)}$$

presque partout sur (a, b). Mais cette inégalité est évidente, car on a, presque partout sur (a, b),

$$\frac{\phi''(x)}{\phi'(x)} = \frac{f''_+(x)}{f'(x)}$$

et, d'autre part, pour chaque $x \in (a, b)$,

$$\frac{f''_+(x)}{f'(x)} \geqslant \frac{f''(x)}{f'(x)}.$$

Le théorème 2 est ainsi complètement démontré.

Remarque. L'hypothèse que $f''(x)$ existe, finie, non seulement presque partout, mais partout sur (a, b), a été posée seulement pour assurer la continuité absolue de $f'(x)$ sur chaque intervalle compact contenu dans (a, b).

LITTÉRATURE

1. G. Szekeres, *On a property of monotone and convex functions*, Proc. Amer. Math. Soc., 7 (1956), 351–3.
2. S. Saks, *Théorie de l'intégrale* (Warszawa, 1933).
3. ——— *Theory of the integral* (Warszawa-Lwow, 1937).

Institut de Matematică
București

SOME RESULTS ON UNIQUENESS AND SUCCESSIVE APPROXIMATIONS

FRED BRAUER

1. In the theory of ordinary differential equations, there is a strange relationship between uniqueness of solutions and convergence of the successive approximations. There are examples of differential equations with unique solutions for which the successive approximations do not converge **(8)** and of differential equations with non-unique solutions for which the successive approximations do converge **(2)**. However, in spite of the known logical independence of these two properties, almost all conditions which assure uniqueness also imply the convergence of the successive approximations. For example, the hypotheses of Kamke's general uniqueness theorem **(5)**, have been shown by Coddington and Levinson to suffice for the convergence of successive approximations, after the addition of one simple monotonicity condition **(4)**. There is one counterexample to this "principle," a generalization of Kamke's result, to which another condition in addition to a monotonicity assumption must be added before convergence of the successive approximations can be proved **(2)**.

In this paper, we shall prove a pair of theorems, generalizing the results of Kamke, and Coddington and Levinson, and conforming to the "principle" mentioned above. The generalization here is in a different direction from that given in **(2)**, but we shall alsoindicate how the results obtained here can be formulate and proved in the more general setting used there.

The statements of our results involve a pair of conditions, each relating to a first order differential equation. We assume the existence of two functions, controlling, in a sense, the behaviour of the solutions of each one of these equations, and a relation between the growths of these functions near the origin. The complexity of the results is compensated for by the large number of special cases which can be obtained by appropriate choice of these functions.

2. We consider the initial value problem

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}), \qquad \mathbf{x}(0) = 0, \tag{1}$$

where $\mathbf{x}$ and $\mathbf{f}$ are n-dimensional vectors. As usual, the norm $|\mathbf{x}|$ of a vector $\mathbf{x}$ will denote the sum of the absolute values of the components of $\mathbf{x}$ **(3)**. Let $\psi_i(t, r)$ $(i = 1, 2)$ be continuous non-negative functions defined for $0 < t < a$,

Received November 6, 1958.

$r \geqslant 0$, which are monotone non-decreasing in r for each fixed t. We will always assume that $\mathbf{f}$ satisfies the pair of conditions

$$|\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)| \leqslant \psi_i(t, |\mathbf{x}_1 - \mathbf{x}_2|), \qquad (i = 1, 2), \tag{2}$$

for $(t, \mathbf{x}_1)$ and $(t, \mathbf{x}_2)$ in a region $0 < t < a$, $|\mathbf{x}| < b$.

THEOREM 1. *Let $\mathbf{f}(t, \mathbf{x})$ be continuous and satisfy (2) in a region $0 < t < a$, $|\mathbf{x}| < b$. Suppose $A(t)$ and $B(t)$ are functions on $0 \leqslant t < a$, with $A(0) = B(0) = 0$, such that*

$$\lim_{t \to 0} A(t)/B(t) = 0. \tag{3}$$

Suppose also that all solutions $u(t)$ of

$$u' = \psi_1(t, u) \tag{4}$$

with $u(0) = 0$ obey $u(t) \leqslant A(t)$ on $0 \leqslant t < a$, and that the only solution $v(t)$ of

$$v' = \psi_2(t, v) \tag{5}$$

on $0 \leqslant t < a$ such that

$$\lim_{t \to 0} v(t)/B(t) = 0 \tag{6}$$

is the trivial solution. Then there is at most one solution of (1) on $0 \leqslant t < a$.

Proof. Suppose there are two solutions $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ of (1) on $0 \leqslant t < a$, and let $m(t) = |\mathbf{x}_1(t) - \mathbf{x}_2(t)|$. Then

$$|m'(t)| \leqslant |\mathbf{x}_1'(t) - \mathbf{x}_2'(t)| \leqslant |\mathbf{f}(t, \mathbf{x}_1(t)) - \mathbf{f}(t, \mathbf{x}_2(t))| \leqslant \psi_1(t, m(t)), \tag{7}$$

using (2). Suppose there exists σ, $0 < \sigma < a$, such that

$$m(\sigma) > A(\sigma). \tag{8}$$

Then there is a solution $u(t)$ of (4) passing through the point $(\sigma, m(\sigma))$ and existing on some interval to the left of σ. As far to the left of σ as $u(t)$ exists, it satisfies

$$u(t) \leqslant m(t). \tag{9}$$

To prove (9), we observe that

$$u' = \psi_1(t, u) + \epsilon, u(\sigma) = m(\sigma), \tag{10}$$

has solutions $u(t, \epsilon)$ for all sufficiently small $\epsilon > 0$, existing as far to the left of σ as $u(t)$ exists, and $\lim_{\epsilon \to 0+} u(t, \epsilon) = u(t)$ (**5**, p. 83). Thus it suffices to prove

$$u(t, \epsilon) \leqslant m(t), \tag{11}$$

for all $\epsilon > 0$ and all solutions $u(t, \epsilon)$ of (10). If this inequality does not hold,

there is a least upper bound ζ of numbers $t \leqslant \sigma$ for which (11) is false. Since $m(\sigma) = u(\sigma) = u(\sigma, \epsilon)$ and the functions $m(t)$ and $u(t, \epsilon)$ are continuous,

$$m(\zeta) = u(\zeta, \epsilon), \qquad m'(\zeta) \geqslant u'(\zeta, \epsilon). \tag{12}$$

Then

$$\psi_1(\zeta, m(\zeta)) + \epsilon = \psi_1(\zeta, u(\zeta, \epsilon)) + \epsilon = u'(\zeta, \epsilon) \leqslant m'(\zeta) \leqslant \psi_1(\zeta, m(\zeta)),$$

using (7), (10), and (12). This contradiction proves (11), which, as we have remarked, implies (9).

The solution $u(t)$ can be continued to $t = 0$. If $u(c) = 0$ for some c, $0 < c < \sigma$, we can effect the continuation by defining $u(t) = 0$ for $0 < t < c$; otherwise (9) ensures the possibility of the continuation. Since $m(0) = 0$, $\lim_{t \to 0} u(t) = 0$, and we define $u(0) = 0$. Now we have a solution $u(t)$ of (4) with $u(0) = 0$, and by hypothesis $u(t) \leqslant A(t)$. However, in view of (8) and the definition of $u(t)$, we have $u(\sigma) > A(\sigma)$, a contradiction which proves

$$m(t) \leqslant A(t), \qquad 0 \leqslant t < a. \tag{13}$$

To prove the uniqueness of solutions, we must show that $m(t)$ vanishes identically on $0 \leqslant t < a$. To complete the proof, we must show that this is implied by (13). Proceeding as before but using ψ_2 in place of ψ_1, we obtain $|m'(t)| \leqslant \psi_2(t, m(t))$. The assumption that $m(\tau) > 0$ for some τ, $0 < \tau < a$, yields, by the same argument as before, a solution $v(t)$ of (5) on $0 \leqslant t \leqslant \tau$ such that $v(\tau) = m(\tau)$, $0 \leqslant v(t) \leqslant m(t)$, $v(0) = 0$. Then

$$0 \leqslant \lim_{t \to 0} v(t)/B(t) \leqslant \lim_{t \to 0} m(t)/B(t) \leqslant \lim_{t \to 0} A(t)/B(t) = 0,$$

using (3) and (13). But by hypothesis this implies that $v(t)$ is identically zero, which contradicts $v(\tau) = m(\tau) > 0$, and therefore $m(t)$ vanishes identically on $0 \leqslant t < a$, which completes the proof of the theorem.

3. The successive approximations to the solution of (1) are defined by

$$\mathbf{x}_0(t) = 0, \mathbf{x}_{j+1}(t) = \int_0^t \mathbf{f}(s, \mathbf{x}_j(s))ds, \qquad (j = 0, 1, \ldots). \tag{14}$$

THEOREM 2. *Let* $\mathbf{f}(t, \mathbf{x})$ *be continuous in a region* $0 < t < a$, $|\mathbf{x}| \leqslant b$ *and bounded in norm by* M *in this region. If the hypotheses of Theorem* 1 *are satisfied then the successive approximations* (14) *converge uniformly on the interval* $0 \leqslant t < \alpha$, *where* $\alpha = \min(a, b/M)$, *to the unique solution of* (1).

Proof. It follows easily from the definition (14) of the successive approximations that they satisfy $|\mathbf{x}_j(t_1) - \mathbf{x}_j(t_2)| \leqslant M|t_1 - t_2|$ in the interval $0 \leqslant t < \alpha$. This implies that the sequence $\{\mathbf{x}_j(t)\}$ is equicontinuous on this interval. Taking $t_2 = 0$, we have $|\mathbf{x}_j(t_1)| \leqslant Mt_1 \leqslant b$, and thus the sequence is also uniformly bounded. Therefore there exists a subsequence $\{\mathbf{x}_{jk}(t)\}$ which

converges uniformly to a function $\mathbf{x}(t)$ on $0 \leqslant t < \alpha$. Because of (14) and the continuity of $\mathbf{f}$, the sequence $\{\mathbf{x}_{jk+1}(t)\}$ converges uniformly to a function

$$\mathbf{x}^*(t) = \int_0^t \mathbf{f}(s, \mathbf{x}(s))ds.$$

We will prove that $\mathbf{x}_{j+1}(t) - \mathbf{x}_j(t) \to 0$ as $j \to \infty$. This will imply, because of (14), that $\mathbf{x}(t) = \mathbf{x}^*(t)$, so that $\mathbf{x}(t)$ is a solution of (1). Since the solution of (1) is unique, every convergent subsequence of $\{\mathbf{x}_j(t)\}$ converges to a solution, hence to the same solution, which shows that the original sequence converges to $\mathbf{x}(t)$ on $0 \leqslant t < \alpha$. Because of the equicontinuity of the sequence, this convergence is uniform.

We define

$$\mathbf{w}_j(t) = \mathbf{x}_{j+1}(t) - \mathbf{x}_j(t), \qquad m(t) = \limsup_{j \to \infty} |\mathbf{w}_j(t)|.$$

Then $m(0) = 0$, and $m(t)$ is continuous on $0 \leqslant t < \alpha$, since it is the upper limit of a uniformly bounded equicontinuous sequence of functions. We must show that $m(t)$ vanishes identically on $0 \leqslant t < \alpha$. Using (14) and (2), we obtain

$$|\mathbf{w}_{j+1}(t+h) - \mathbf{w}_{j+1}(t)| \leqslant \int_t^{t+h} |\mathbf{f}(s, \mathbf{x}_{j+1}(s)) - \mathbf{f}(s, \mathbf{x}_j(s))|ds \tag{15}$$

$$\leqslant \int_t^{t+h} \psi_1(s, |\mathbf{w}_j(s)|)ds.$$

Because of the continuity of m and the equicontinuity of $\{\mathbf{w}_j\}$, given any $\epsilon > 0$, there exists an integer $N(\epsilon)$, independent of s and j, such that

$$|\mathbf{w}_j(s)| < m(s) + \epsilon, \qquad j > N(\epsilon). \tag{16}$$

It follows from (15) and (16) that

$$|\mathbf{w}_{j+1}(t+h) - \mathbf{w}_{j+1}(t)| \leqslant \int_t^{t+h} \psi_1(s, m(s)+\epsilon)ds, \quad j > N(\epsilon). \tag{17}$$

It follows easily from the definition of m that

$$|m(t+h) - m(t)| \leqslant \limsup_{j \to \infty} |\mathbf{w}_{j+1}(t+h) - \mathbf{w}_{j+1}(t)|.$$

Combining this with (17) and then letting $\epsilon \to 0$, we obtain

$$|m(t+h) - m(t)| \leqslant \int_t^{t+h} \psi_1(s, m(s))ds, \tag{18}$$

using the continuity and monotonicity of $\psi_1(t, r)$ in r. The inequality (18) implies that $m'(t)$ exists on any interval $(t, t+h)$, for $t \geqslant 0$, and that $|m'(t)| \leqslant \psi_1(t, m(t))$. The argument used in the proof of Theorem 1, beginning with (7), proves

$$m(t) \leqslant A(t). \tag{19}$$

To complete the proof, we must show that (19) implies $m(t) \equiv 0$. The argument is much the same as the last stage of the proof of Theorem 1. Suppose $m(\tau) > 0$ for some τ, $0 < \tau < a$. A repetition of the first part of the proof, using ψ_2 in place of ψ_1, gives a solution $v(t)$ of (5) on $0 \leqslant t \leqslant \tau$ such that $v(\tau) = m(\tau)$, $0 \leqslant v(t) \leqslant m(t)$, $v(0) = 0$. Then $0 \leqslant \lim_{t\to 0} v(t)/B(t) \leqslant \lim_{t\to 0} m(t)/B(t) \leqslant \lim_{t\to 0} A(t)/B(t) = 0$, using (3) and (19). By hypothesis, this implies that $v(t)$ vanishes identically, contradicting $v(\tau) = m(\tau) > 0$, and it follows that $m(t) = 0$, $0 \leqslant t < a$.

4. The results of this paper were originally suggested by consideration of a function $\mathbf{f}(t, \mathbf{x})$ satisfying the pair of conditions

$$|\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)| \leqslant C|\mathbf{x}_1 - \mathbf{x}_2|^{\alpha}, \tag{20}$$

$$|\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)| \leqslant k|\mathbf{x}_1 - \mathbf{x}_2|/t, \tag{21}$$

where C, α, and k are constants with $0 < k$, $0 < \alpha < 1$, $k(1 - \alpha) < 1$. It was shown by Krein and Krasnosel'skii **(6)** that this pair of conditions implies the uniqueness of solutions of (1), and by Luxemburg **(7)** that the successive approximations (14) converge. In **(1)** these results were proved by methods analogous to those used here. They are special cases of Theorems 1 and 2 obtained by taking $\psi_1(t, r) = Cr^{\alpha}$, $\psi_2(t, r) = kr/t$. It is easily verified that the hypotheses of Theorems 1 and 2 are satisfied, with $A(t) = C(1 - \alpha)t^{1/(1-\alpha)}$, $B(t) = t^k$, and the condition $k(1 - \alpha) < 1$ gives (3). Thus the results of Krein–Krasnosel'skiĭ and Luxemburg are contained in those of this paper.

The general uniqueness theorem of Kamke **(5)** and the successive approximation theorem of Coddington and Levinson **(4)** are also contained in this paper. In fact, the proofs used here are quite similar to the proofs of these results **(3**, chapter 2**)**. The central hypothesis there is the existence of a function $\psi(t, r)$, continuous and non-negative in $0 < t < a$, $r \geqslant 0$, and monotone non-decreasing in r for each fixed t, such that there is no non-trivial solution of

$$r' = \psi(t, r), \qquad r(0) = r'(0) = 0, \tag{22}$$

on $0 \leqslant t < a$, and such that

$$|\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)| \leqslant \psi(t, |\mathbf{x}_1 - \mathbf{x}_2|),$$

for $(t, \mathbf{x}_1)$, $(t, \mathbf{x}_2)$ in a region $0 < t < a$, $|\mathbf{x}| \leqslant b$. That this result is contained in the present results is seen by taking $\psi_1(t, r) = \psi(t, r)$ and $\psi_2(t, r) = 2M$, where M is a bound for $|\mathbf{f}(t, \mathbf{x})|$ in $0 < t < a$, $|\mathbf{x}| < b$. Then we can take $B(t) = t$. If $A(t)$ is a solution of (4) with $A(0) = A'(0) = 0$, the condition (3) follows from

$$0 = A'(0) = \lim_{t\to 0} A(t)/t = \lim_{t\to 0} A(t)/B(t),$$

and the hypotheses of Theorems 1 and 2 are satisfied.

5. Now we obtain a generalization of the Krein–Krasnosel'skiĭ–Luxemburg results by application of our theorems. This will illustrate the flexibility of the method in generating new results. We retain the condition (21), corresponding to $\psi_2(t, r) = kr/t$, $B(t) = t^k$, but replace (20) by

$$|\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)| \leqslant C|\mathbf{x}_1 - \mathbf{x}_2|^\alpha t^\beta, \qquad \alpha < 1, \beta > -1, \tag{23}$$

corresponding to the choice $\psi_1(t, r) = Cr^\alpha t^\beta$. It is easy to solve (4) with this choice of ψ_1, and we find that we can take $A(t) = pt^{(\beta+1)/(1-\alpha)}$, where p is a constant depending only on α, β, C. For (3) to be satisfied, we must have $(\beta + 1)/(1 - \alpha) - k > 0$, or $\beta + 1 > k(1 - \alpha)$. The results of **(1; 6; 7)** are the special case $\beta = 0$. If $(\beta + 1)/(1 - \alpha) \leqslant 1$, the condition (23) alone would suffice, as can be seen from the Kamke-Coddington-Levinson results. Thus we assume $\beta + 1 > 1 - \alpha$, or $\beta > -\alpha$.

Slightly more generally, we can replace the condition (23) by

$$|\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)| \leqslant C|\mathbf{x}_1 - \mathbf{x}_2|^\alpha \lambda(t), \tag{24}$$

where $\alpha < 1$ and $\lambda(t) = o[t^{k(1-\alpha)-1}]$ as $t \to 0$. This corresponds to the choice $\psi_1(t, r) = Cr^\alpha \lambda(t)$, which yields

$$[A(t)]^{1-\alpha} = p \int_0^t \lambda(t)dt.$$

If $\lambda(t)$ satisfies the above growth condition, we have

$$\int_0^t \lambda(t)dt = o[t^{k(1-\alpha)}],$$

and $A(t) = o(t^k) = o(B(t))$, so that the hypotheses of our theorems are satisfied.

These examples illustrate suitable choices of $\psi_1(t, r)$ and $A(t)$. There is less flexibility in the choice of $\psi_2(t, r)$, as the definition of $B(t)$ implies that there can be no non-trivial solution of (5) which vanishes on an interval near the origin.

6. The results obtained in this paper can also be given in the more general setting of **(2)**. As the proofs are quite similar to those given here, with the same alterations used in **(2)**, we shall only outline the results, without proofs. Instead of using a norm $|\mathbf{x}|$ for vectors $\mathbf{x}$, we use a function $V(t, \mathbf{x})$ defined for real t and vectors $\mathbf{x}$ with non-negative real values, which is continuous in $(t, \mathbf{x})$, has one-sided partial derivatives with respect to t and the components of $\mathbf{x}$, and whose vanishing implies $\mathbf{x} = 0$. We use V_t to denote a partial derivative of V with respect to t, V_x to denote some gradient vector of V, and . to denote the usual scalar product of vectors. Any condition which involves V_t or V_x is understood to be required for all one-sided derivatives.

The only change in the hypotheses of the theorems is that the pair of conditions (2) are replaced by

$$V_t(t, \mathbf{x}_1 - \mathbf{x}_2) + V_x \,.\, (\mathbf{f}(t, \mathbf{x}_1) - \mathbf{f}(t, \mathbf{x}_2)) \leqslant \psi_i(t, V(t, \mathbf{x}_1 - \mathbf{x}_2)), \; i = 1, 2,$$

in the generalization of Theorem 1. For the generalization of Theorem 2, we must add the condition

$$V(t, \int_{t-h}^{t} [\mathbf{f}(s, \mathbf{x}_1(s)) - \mathbf{f}(s, \mathbf{x}_2(s))]ds)$$

$$< \int_{t-h}^{t} \{V_t(s, \mathbf{x}_1(s) - \mathbf{x}_2(s)) + V_x \cdot [\mathbf{f}(s, \mathbf{x}_1(s)) - \mathbf{f}(s, \mathbf{x}_2(s))]\}\, ds,$$

for any continuous $\mathbf{f}$, $\mathbf{x}_1$, $\mathbf{x}_2$, and the condition $V(\max(t_1, t_2), \mathbf{x}_1 + \mathbf{x}_2) \leqslant V(\mathbf{x}_1, t_1) + V(\mathbf{x}_2, t_2)$. This condition is essentially just a weakened form of the triangle inequality for the function $V(t, \mathbf{x})$. The other hypotheses of Theorems 1 and 2 and the conclusions remain unchanged.

REFERENCES

1. F. Brauer, *A note on uniqueness and convergence of successive approximations*, Can. Math. Bulletin *2* (1959), 5–8.
2. F. Brauer and S. Sternberg, *Local uniqueness, existence in the large, and the convergence of successive approximations*, Amer. J. Math., *80* (1958), 421–30.
3. E. A. Coddington and Norman Levinson, *Theory of ordinary differential equations* (New York, 1955).
4. ——— *Uniqueness and the convergence of successive approximations*, J. Indian Math. Soc., *16* (1952), 75–81.
5. E. A. Kamke, *Differentialgleichungen reeller Funktionen* (Leipzig, 1930).
6. M. A. Krasnosel'skii and S. G. Krein, *On a class of uniqueness theorems for the equation* $y' = f(x, y)$, Uspehi Mat. Nauk (N.S.) *11* (1956), 209–13 (Russian).
7. W. A. J. Luxemburg, *On the convergence of successive approximations in the theory of ordinary differential equations*, Can. Math. Bulletin, *1* (1958), 9–20.
8. M. Müller, *Ueber das Fundamentaltheorem in der Theorie der gewöhnlichen Differentialgleichungen*, Math. Zeit., *26* (1927), 619–45.

University of British Columbia

THE ASYMPTOTIC BEHAVIOUR OF THE LAURENT COEFFICIENTS

MAX WYMAN

Introduction. Let $G(z)$ be a function of a complex variable, regular in the annulus $0 \leqslant a < |z| < b \leqslant \infty$. We shall assume there exists a curve within the annulus for which

$$\lim_{|z| \to b} G(z) = \infty,$$

provided z is restricted to be a point of this curve. Under these restrictions $G(z)$ has a Laurent expansion of the form

$$G(z) = \sum_{-\infty}^{\infty} a_n z^n, \tag{1.1}$$

where the Laurent coefficients a_n have the integral representation

$$a_n = (2\pi i)^{-1} \int_C G(z)\, z^{-(n+1)}\, dz, \tag{1.2}$$

and C can be any contour, within the domain of regularity, that encloses $z = 0$. We shall also assume that the a_n are all real numbers. Using the usual complex conjugate notation, we can, therefore, write

$$\overline{G(z)} = G(\bar{z}). \tag{1.3}$$

The problem of determining the asymptotic behaviour of a_n as $n \to \infty$ is very old in mathematical literature and appears in many forms and disguises. It has been solved for specific classes of functions by many people using a multitude of methods. For certain lacunary type series a_n can have an almost chaotic behaviour for large n. It is, therefore, too much to expect that there exists a single method that will give the asymptotic behaviour of a_n for the class of all functions that possess Laurent expansions. We shall, therefore, in this paper, consider only a single method that yields the asymptotic behaviour of a_n for a particular class of generating functions $G(z)$ which will be called admissible. The major goal is to make the class of admissible functions as large as possible.

As far as we are aware the first attack on problems of this type, that could claim any degree of generality, is due to Darboux **(1)**. His class of admissible functions possessed Maclaurin expansions with a finite radius of convergence. On the circle of convergence $G(z)$ was allowed to have only a finite number of singularities of a particular type. The success of Darboux's method is almost

Received July 9, 1958.

entirely due to the fact that his assumptions allow us to deform the contour C beyond the circle of convergence and ultimately to prove that only small contours surrounding the singularities of $G(z)$ contribute to the asymptotic behaviour of a_n. This procedure is almost characteristic of all methods that depend for their success on the classification of the number and types of singularities the generating function $G(z)$ is allowed to have on the circle of convergence.

Such deformations are not available if the circle of convergence is a natural boundary of $G(z)$. In particular the partition function

$$G(z) = \prod_{n=1}^{\infty} (1 - z^n)^{-1} = \sum_{n=0}^{\infty} p(n)\, z^n \tag{1.4}$$

falls into this classification. There exists, in the literature, several derivations of the behaviour of $p(n)$ as $n \to \infty$.

The success of known methods of studying the asymptotic behaviour of sequences, defined by integrals, lies almost entirely in the fact that the behaviour is determined by a knowledge of $G(z)$ in the neighbourhoods of certain points called critical points. Erdélyi **(2)** has pointed out that there does not exist any general theory of critical points and that only a few types of integrals have been studied. We shall provide such a theory for integrals of the type (1.2) and for which $G(z)$ belongs to our class of admissible functions.

Among recent papers dealing with problems of this type we note those of Szekeres **(8)**, Hayman **(4)**, and Moser and Wyman **(5; 6)**. Szekeres finds a complete asymptotic expansion for a specific example. Hayman introduced the idea of a class of admissible functions and found the first term of an asymptotic expansion for the Maclaurin coefficient of all members of his class of functions. Wyman and Moser find complete asymptotic expansions for the Maclaurin coefficients of a class of integral functions. We note these papers in particular because in each case there is a more or less common pattern of attack. In our present paper we shall attempt to generalize the general pattern of attack of the papers mentioned above.

2. Asymptotic expansions. In applying a method of critical points to integrals of the form (1.2) we place suitable restrictions on $G(z)$ that enable us to determine the asymptotic behaviour of a_n by considering only restricted portions of the contour C. It is thus clear that such a method implies a criterion by means of which we may recognize the portions of C that are to be retained and to recognize also the portions of C that may be discarded. This criterion is found in the choice of a suitable definition of an asymptotic expansion.

Since Poincaré's introduction of a definition of an asymptotic expansion, the concept has been generalized in many ways. Erdélyi **(2)** has given an elegant treatment of the concept which can be used to develop a general theory of asymptotic expansions. However, his treatment imposes certain

restrictions that are neither necessary nor desirable for the type of problem we shall consider in this paper. In a private conversation Professor Erdélyi has given me a more general definition that completely fills our present needs. We shall, however, use only a very specialized version of Erdélyi's later definition.

Throughout the paper r shall denote a real variable lying in the fixed interval I given by $0 \leqslant a \leqslant r < b \leqslant \infty$. We shall be interested in the behaviour of certain real or complex valued functions $F(r)$ as $r \to b$. The end-point a plays no essential role in our discussion except in so far as all statements are required to have meaning only if r is restricted to be a value in I.

Definition 2.1. A real valued positive function $V(r)$ is called a *comparison function* if

$$\lim_{r \to b} V(r) = \infty.$$

Definition 2.2. Two real or complex valued functions $f(r)$ and $g(r)$ are said to be *asymptotically equal* at $r = b$, with respect to the comparison function $V(r)$, if

$$\lim_{r \to b} V^n (f(r) - g(r)) = 0 \tag{2.1}$$

for every non-negative integer n.

We shall write this relationship as

$$f(r) = g(r) \tag{2.2}$$

or in terms of the order relation o as

$$f(r) - g(r) = o(V^{-n}), \tag{2.3}$$

for every non-negative integer n.

Since we shall usually require b and $V(r)$ to remain fixed in our discussion we shall often delete the qualifying phrases in the definition of asymptotic equality.

If there exists a value r_0 in I and positive real numbers α, β such that

$$|f(r) - g(r)| = O(\exp(-\beta V^{\alpha})) \tag{2.4}$$

for all $r \geqslant r_0$, then $f(r) = g(r)$.

Asymptotic equality is an equivalence relation and divides the class of complex functions defined on I into equivalence classes. In our development we shall usually replace a function $f(r) = 0$ by zero. If $b = \infty$, $V(r) = \log r$ then r, $(r^2 + 1)/r$ are examples of a pair of asymptotically equal functions.

Definition 2.3. Let $A_k(r)$, $k = 0, 1, 2, \ldots$ be an infinite sequence of complex valued functions. The formal series

$$\sum_{k=0}^{\infty} A_k V^{-k}$$

is called an *asymptotic expansion* of a function $F(r)$ if, for every non-negative integer m,

$$\lim_{r\to b} V^m \left(F - \sum_{k=0}^{m} A_k V^{-k}\right) = 0. \tag{2.5}$$

We write

$$F \sim \sum_{k=0}^{\infty} A_k V^{-k} \tag{2.6}$$

or

$$F - \sum_{k=0}^{m} A_k V^{-k} = o(V^{-m}). \tag{2.7}$$

The asymptotic equality of two functions is a special case of an asymptotic expansion with $A_k = 0$, $k > 1$. It is also easy to show that the definition of an asymptotic expansion implies $A_k = o(V)$, $k > 1$ and that the asymptotic expansion of a function is not unique. Any of the $A_k(r)$ may be replaced by an asymptotically equal function.

Definition 2.4. If $F(r)$ and $H(r)$ are two complex valued functions such that

$$F/H \sim \sum_{k=0}^{\infty} A_k V^{-k} \tag{2.8}$$

then we write

$$F \sim H \sum_{k=0}^{\infty} A_k V^{-k}. \tag{2.9}$$

At first glance it might be thought that Definition (2.4) is redundant in the light of the fact that the $A_k(r)$ of Definition (2.3) are allowed to be functions of r. This, however, is not so because (2.9) implies only that $A_k = o(V)$ while

$$F \sim \sum_{k=0}^{\infty} (H A_k) V^{-k} \tag{2.10}$$

implies $HA_k = o(V)$. Hence (2.10) may or may not place a restriction on H that is not implied by (2.9).

The major distinction between our use at present of the meaning of an asymptotic expansion and Erdélyi's published version is the fact that we do not require $A_k V^{-k}$ to be an asymptotic sequence, even though V^{-k} is such a sequence. This situation could exist, for example, if some of the $A_k = 0$.

3. Relevant paths. Returning to (1.1), we may, without any loss of generality, assume that we are interested in the asymptotic behaviour of a_n for large positive n. The behaviour of a_n, for large negative n, could then be obtained from the generating function $G(1/z)$. If $G(z)$ has the form

$$G(z) = P_m(z) + Q(z^{-1}) \tag{3.1}$$

where $P_m(z)$ is a polynomial of degree m and Q is a regular function of its argument, then our problem is trivial. For such a function $a_n = 0, n > m + 1$. Hence we shall exclude all functions of the form (3.1) from the remainder of our discussion.

Using the polar form $z = r\exp(i\theta)$ we define $M(r, \theta)$ by

$$M(r, \theta) = |G(r\exp(i\theta))|. \tag{3.2}$$

Since we shall always operate within the domain of regularity of $G(z)$ for which (1.1) is true, r shall always lie in the interval I of the previous section. The function $M(r, \theta)$ is a periodic function of θ with a period of 2π and has continuous partial derivatives of all orders except possibly at the zeros of $G(z)$. The function $M(r)$ defined by

$$M(r) = \sup_{|z|=r} M(r, \theta) \tag{3.3}$$

is the so-called maximum modulus function and it is known that there always exists a point $z = z(r)$, on $|z| = r$, such that $|G(z(r))| = M(r)$. By varying r, $z = z(r)$ can be considered as the parametric representation of a curve in the complex z plane which is called a path of maximum modulus. The path of maximum modulus need not be unique. From the assumptions made, in the Introduction, on $G(z)$ it is clear that

$$\lim_{r\to b} M(r) = \infty$$

and hence $\log M(r)$ is a comparison function. The choice of $\log M(r)$ as a comparison function may, however, introduce a complexity that is neither necessary nor desirable. Hardy (3) has shown that $G(z) = \exp(\sin z)\exp(\exp z^2)$ has $M(r) = \exp|\sin r|\exp(\exp r^2)$. The corresponding path of maximum modulus is $\theta = 0, 2k\pi \leqslant r \leqslant (2k+1)\pi$ and $\theta = \pi, (2k+1)\pi \leqslant r \leqslant (2k+2)\pi$. Hence the path of maximum modulus has an infinite number of discontinuities. For this reason, we shall see, the *a priori* choice of $\log M(r)$ as a comparison function often introduces a discontinuous picture of our problem that can easily be avoided. Previous authors have featured $\log M(r)$ in their discussions and have required $\theta = 0$ to be the continuous path of maximum modulus. Such a starting point greatly reduces, for no valid reason, the extent of our class of admissible generating functions $G(z)$.

Let us return to (3.2) and consider the stationary values of $M(r, \theta)$ which are given as solutions of the equation

$$\frac{\partial M}{\partial \theta} = 0. \tag{3.4}$$

There always exists at least one real solution of (3.4) $\theta = \theta(r)$. In fact our assumption that all a_n are real implies that $\theta = 0$ or $\theta = \pi$ are always solutions of (3.4) except possibly for the case when $M(r, 0) = 0$ or $M(r, \pi) = 0$. Implicit function theory tells us that if (r_0, θ_0) is a solution of (3.4), for which

$$\frac{\partial^2 M}{\partial \theta^2} \neq 0,$$

then there always exist neighbourhoods $|r - r_0| < \delta$, $|\theta - \theta_0| < \delta$ for which the solution $\theta = \theta(r)$ is unique and has a continuous derivative. We may, therefore, consider such a solution $\theta = \theta(r)$ as the polar equation of a path, in the complex z plane, which lies within the domain of regularity of $G(z)$. If we traverse such a path, with increasing r, the path remains unique until we strike a point at which

$$\frac{\partial^2 M}{\partial \theta^2} = 0.$$

At such a point the path may cease to be unique and several branches may appear. However if r is sufficiently close to b, then (9, pp. 22-7), the number of such branch paths remains finite. Although $\theta(r)$ need not be differentiable at a branch point, the continuity of $\theta(r)$ is preserved. All paths determined as solutions of (3.4) shall be called the stationary paths of $G(z)$.

We interrupt our general development to consider the specific function $\exp(z/1 + z^2))$ which gives considerable insight into the general situation. The solutions of (3.4) are

$$\theta = 0,\ \theta = \pi,\ 0 \leqslant r \leqslant \sqrt{2} - 1 \tag{3.5}$$

$$\theta = 0,\ \theta = \pi,\ \theta = \pm \arccos((1 - r^2)/2r),\ \theta = \pi \pm \arccos((1 - r^2)/2r),\ \sqrt{2} - 1 \leqslant r \leqslant 1. \tag{3.6}$$

The points, $(0, \sqrt{2} - 1)$, $(\pi, \sqrt{2} - 1)$, are branch points and there exist six stationary paths by means of which we may leave $z = 0$ and arrive at the boundary $|z| = 1$.

We note that the number of intersections of the stationary paths with the circle $|z| = r$ is a function of r and indicates the possibility that, for other generating functions $G(z)$, that the number of such intersections may tend to ∞ as $r \to b$. Let us now consider the stationary path $\theta = 0$, $0 \leqslant r \leqslant 1$. For a fixed value of r and variable θ, $M(r, 0)$ is a maximum of $M(r, \theta)$, for $0 \leqslant r < \sqrt{1} - 2$, and then $M(r, 0)$ changes its character to become a minimum of $M(r, \theta)$ for $\sqrt{2} - 1 < r < 1$. Again the example indicates the possibility that there may exist stationary paths $\theta = \theta(r)$ such that, along such a path, $M(r, \theta(r))$ may change its character from being a maximum of $M(r, \theta)$ to being a minimum of $M(r, \theta)$ an infinite number of times.

Returning to the general case we will denote by L_k, $(k = 1, 2, \ldots, T)$, the stationary paths contained in the annulus $r_0 \leqslant |z| \leqslant r < b$. We note that the integer T may be a function of r. The corresponding polar equations are written as $\theta = \theta_k(r)$ and $M_k(r)$ is defined by

$$M_k(r) = M(r, \theta_k(r)). \tag{3.7}$$

Assumption (1). For r sufficiently close to b there exists a continuous stationary path, with polar equation $\theta = \theta_1(r)$, by means of which we can reach the boundary $r = b$. Further $M(r, \theta)/M_1$ is bounded uniformly in θ.

Since the maximum modulus $M(r)$ is always attained for every value of r

at some value of θ assumption (1) implies that $M(r)/M_1$ is bounded as $r \to b$. However we have previously assumed that

$$\lim_{r \to b} M(r) = \infty.$$

This means that

$$\lim_{r \to b} M_1(r) = \infty.$$

Thus $\log M_1(r)$ is a comparison function by means of which we may give meaning to the concepts of asymptotic equality and asymptotic expansions. Throughout the remainder of the paper $\log M_1(r)$ will be our comparison function and $(\log M_1(r))^{-m}$ will be our asymptotic sequence.

Definition. A stationary path for which $M_k/M_1 = 0$ will be called an *irrelevant path.* All other stationary paths will be called *relevant paths.*

Assumption (2). We shall assume that the relevant paths can be identified in such a way that the following properties are true for r sufficiently close to b.

(a) The relevant paths are all continuous curves by means of which we may reach the boundary $r = b$.

(b) If $\theta = \theta_k(r)$ is a relevant path then a constant $K_k > 0$ and a non-negative integer m exists such that $M_k/M_1 \geqslant K_k (\log M_1)^{-m}$.

(c) The number N of relevant paths is fixed and independent of r.

(d) For every relevant path $\theta = \theta_k(r)$ the

$$\lim_{r \to b} \theta_k(r)$$

exists. Further if $\theta = \theta_k(r)$ and $\theta = \theta_s(r)$ are distinct relevant paths then

$$\lim_{r \to b} \theta_k(r) \neq \lim_{r \to b} \theta_s(r).$$

(e) Along every relevant path the function $M(r, \theta)$ has the property

$$\frac{\partial^2 M}{\partial \theta^2} < 0.$$

An effect of assumption (2) is that for r close to b the relevant paths can have no point of intersection. Such a point would be a branch point at which

$$\frac{\partial^2 M}{\partial \theta^2} = 0.$$

This would of course contradict the assumption. The assumption also guarantees that there can be no point of intersection of relevant paths even at $r = b$. We may also conclude that the number of intersections of the relevant paths with the circle $|z| = r < b$ is equal to the fixed number N. Hence without any loss of generality we may assume that the relevant paths are L_k, $(k = 1, 2, \ldots, N)$ and that they have been numbered in a counter-clockwise direction beginning

with the featured path L_1. Along every relevant path the value $M_k(r)$ is, for fixed r and variable θ, a maximum value of $M(r, \theta)$ as $r \to b$. Finally, we also note that the functions $\theta_k(r)$ corresponding to relevant paths all have continuous first derivatives.

Since (any constant)$/M_1(r) \to 0$ it is clear that $M_k(r)$, $(k = 1, 2, \ldots, N)$ must all be unbounded functions of r as $r \to b$. Since $M(r)/M_1(r)$ and $M_1(r)/M(r)$ are both bounded functions we must have

$$\lim_{r \to b} \log M(r)/\log M_1(r) = 1. \tag{3.8}$$

This result can be extended to

$$\lim_{r \to b} (\log M_k(r)/\log M_1(r)) = 1, \qquad k = 1, 2, \ldots, N. \tag{3.9}$$

It is not difficult to show that the function $\exp(\sin z)\exp(\exp z^2)$ obeys assumptions (1) and (2).

4. Functional behaviour along relevant paths. In the present section we shall investigate the functional behaviour of $\log G(z)$ and its derivatives along relevant paths. Since it turns out to be simpler to study the derivatives with respect to $\log z$ we introduce the operator H by

$$H = z\frac{d}{dz} \tag{4.1}$$

and attach the usual meaning to the iterated operator H^m. Further, the symbol $H^m F(w)$ shall always mean $H^m F(z)]_{z=w}$. In this notation (3.4) becomes

$$\operatorname{Im}(H \log G(z)) = 0 \tag{4.2}$$

and hence the quantity $H \log G(z)$ is always real along every relevant path L_k.

Let us denote the points of intersection of L_k with $|z| = r < b$ by

$$z = z_k(r), \quad |z_k| = r, \quad k = 1, 2, \ldots, N. \tag{4.3}$$

For variable r, (4.3) provides a parametric representation of the relevant paths. Further $z_k(r)$ has a continuous derivative with respect to r as $r \to b$, since $\theta_k(r)$ has this property. Since $|G(z_k)| = M_k(r)$ and $M_k(r)$ is a maximum of $M(r, \theta)$ we must have $G(z_k) \neq 0$. Thus all of the functions $H^m \log G(z_k(r))$ have a continuous derivative with respect to r for r sufficiently close to b. The condition

$$\frac{\partial^2 M}{\partial \theta^2} < 0$$

can be translated to read

$$Rl\,(H^2 \log G(z)) > 0, \tag{4.4}$$

along every relevant path.

LEMMA 4.1. *The function $H \log G(z)$ is increasing in r and has no upper bound along any relevant path.*

Proof. Let z_0 be any fixed point along a relevant path and suppose z is a variable point along the same path. Along such a path $H \log G(z)$ has been shown to be real. If there exists a constant K such that $H \log G(z) < K$ then the identity

$$\log G(z) = \log G(z_0) + \int_{z_0}^{z} H \log G(z)\, d(\log z), \tag{4.5}$$

where the path of integration is the relevant path, implies

$$|G(z)| \leqslant |G(z_0)|\ (r/|z_0|)^K. \tag{4.6}$$

However, since our path of integration is a relevant path L_k we must have $|G(z)| = M_k(r)$. If b is finite, (4.6) implies that $M_k(r)$ is a bounded function of r as $r \to b$. This contradicts our assumption that L_k is a relevant path. If b is infinite the proof is somewhat more difficult but follows directly from known complex function theory as long as we remember that functions of the form (3.1) are excluded.

Along any relevant path we know that $H \log G(z)$ has a continuous first derivative with respect to r, if r is close enough to b. An easy computation gives

$$r \frac{d}{dr} H \log G(z) = Rl\ H^2 \log G(z) \left(1 + r^2 \left(\frac{d\theta}{dr}\right)^2\right). \tag{4.7}$$

From (4.4) we have $d/dr\ H \log G(z) > 0$ along a relevant path and hence $H \log G(z)$ must be an increasing function of r along every such path. Since $H \log G(z)$ is an increasing continuous function of r, which does not have an upper bound, along every relevant path we must have

$$\lim_{r \to b} H \log G(z) = +\infty. \tag{4.8}$$

Thus far in our paper we have been discussing the functional behaviour of $G(z)$ and have not related this behaviour to the Laurent coefficients a_n. In order to make a start on obtaining this relationship we shall for the moment allow n to be a large positive real number. Ultimately n will assume its integer meaning in a_n. The results contained in Lemma 4.1 imply that, along any relevant path L_k, the equation

$$H \log G(z) = n \tag{4.9}$$

always has a unique solution $z = Z_k(n)$, for n sufficiently large. Considering n as a variable, the equations

$$z = Z_k(n) \tag{4.10}$$

provide a second parametric representation of L_k. Since it becomes convenient to use both of the parametric representations (4.3) and (4.10), simul-

taneously in some discussions, we introduce the following convention to avoid confusion.

Convention. If r is taken to be an independent variable then n shall mean the function $n = n(r)$ defined by

$$n(r) = H \log G(z_1(r)). \tag{4.11}$$

Thus $z_1(r) = Z_1(n)$ and we may consider both sets of points $z_k(r)$, $Z_k(n)$ as completely determined.

If n is taken to be an independent variable then r is taken to mean the function $r(n)$ defined by

$$r(n) = Z_1(n), \tag{4.12}$$

and again $z_1(r) = Z_1(n)$. As before $z_k(r)$, $Z_k(n)$ are completely determined by this convention. This notation implies

$$\lim_{r \to b} n(r) = \infty, \quad \lim_{n \to \infty} r(n) = b. \tag{4.13}$$

Assumption (3). We shall assume that, for r sufficiently close to b each of the following are true.

(a) There exist positive constants p, P_1, P_2, such that

$$P_1 (\log M_1) \leqslant n(r) \leqslant P_2 ((\log M_1)^{1+p}) \tag{4.14}$$

(b) $Z_k(n)/z_k(r) = 1 + 0(1/n)$.

(c) In the complex w plane there exists a fixed neighbourhood $|w| \leqslant h$ for which the functions $\log G(Z_k(n) \exp(\log M_1 w/n))$ are all regular functions of w. Further,

$$\lim_{r \to b} [\log G(Z_k(n) \exp(\log M_1(r)\, w/n)) - \log G(Z_k(n))]/\log M_1$$

exists, uniformly in w, for all w within and on the boundary of $|w| \leqslant h$. We shall denote this limit by $g_k(w)$ and shall assume $Rlg_k''(0) \neq 0$.

Definition 4.1. Any generating function having a Laurent expansion of the type (1.1) and for which assumptions (1), (2), and (3) are true will be called an admissible function.

The major result of this paper will be that the asymptotic behaviour of the Laurent coefficients of every admissible function can be determined. At this stage the conditions contained in assumption (3) must seem somewhat like pulling rabbits out of a hat. However these conditions do arise quite naturally in the method that we shall use to prove our central theorem.

Part (a) of assumption (3) makes all members of the original class of functions considered by Darboux inadmissible. As such our central theorem will be complimentary to the Darboux result. The second part of the assumption turns out to be sufficient to control the behaviour of the factor $z^{-(n+1)}$ in the integrand of (1.2). I suspect that this part of assumption (3) may be completely unnecessary and that every admissible function may have this

property. The third part of the assumption involves functions that are reminiscent of the indicator functions of integral function theory. In some cases the indicator function can be derived from $g_k(w)$. Obviously the third part of this assumption places some restrictions on the locations of the zeros of $G(z)$.

To be a member of Hayman's class of admissible functions a generating function $G(z)$ has to have a unique path of maximum modulus that is $\theta = 0$. Certainly assumption (1) is considerably weaker than this condition. To this extent our class of admissible functions is considerably more extensive than is Hayman's. It is not, however, possible to say that our class of such functions completely contains Hayman's. Our assumptions have been designed to measure the contribution of any portion of the contour C, of (1.2), to any term in a complete asymptotic expansion. Hayman's assumptions were designed to measure the contributions of C to the first term of such an expansion. If we were to restrict ourselves to Hayman's objective we could, along the lines of the present paper, greatly extend Hayman's class of admissible functions.

When the outer boundary of convergence is finite our assumptions say, in essence, that $G(z)$ may have an infinite number of singularities on $r = b$. However, there can be only a finite number that dominate all the rest. The meaning of dominate is of course derived from our definition of asymptotic equality.

The two specific functions used for illustration in this paper are both admissible functions in our sense and neither is admissible in the Hayman sense. To illustrate another point we shall prove that

$$G(z) = [\exp(1/2(1 - z^2))]/(1 - z) \tag{4.15}$$

is an admissible function. For (4.15) it is not too difficult to prove that

$$z = r \quad \text{and} \quad z = -r \tag{4.16}$$

are the only two possible relevant paths. Since $z = r$ is the path of maximum modulus and since this path is continuous, for this example we choose

$$z_1(r) = r, \qquad z_2(r) = -r. \tag{4.17}$$

Hence

$$M_1(r) = \frac{\exp(\frac{1}{2}(1 - r^2))}{1 - r}, \; M_2(r) = \frac{\exp(\frac{1}{2}(1 - r^2))}{1 + r}. \tag{4.18}$$

Since

$$\lim_{r \to 1} \frac{(\log M_1(r))M_2(r)}{M_1(r)} = \frac{1}{8}, \tag{4.19}$$

we see that both paths are relevant. The fact that

$$\frac{\partial^2 M}{\partial \theta^2} < 0$$

on a relevant path is easily verified. This means assumption (1) is satisfied.

Turning to equation (4.9) the solutions are

$$Z_1(n) = 1 - \frac{1}{2(n)^{1/2}} - \frac{3}{8n} + \frac{1}{8(n)^{3/2}} + \dots, \tag{4.20}$$

$$Z_2(n) = -1 + \frac{1}{2(n)^{1/2}} - \frac{1}{8n} - \frac{1}{8(n)^{3/2}} + \dots. \tag{4.21}$$

Since $Z_1(n) = z_1(r) = r$ we find

$$z_2(r) = -1 + \frac{1}{2(n)^{1/2}} + \frac{3}{8n} - \frac{1}{8(n)^{3/2}} + \dots \tag{4.22}$$

and

$$Z_2(n)/z_2(r) = 1 + 0(1/n). \tag{4.23}$$

Since $\log M_1(r) = \frac{1}{2}(n)^{\frac{1}{2}} + \dots$, the first and second parts of assumption (3) are true. One can readily verify that $g_1(w) = g_2(w) = w/1 - w$ and that the remaining conditions of assumption (3) are true. Thus (4.15) gives a third example of an admissible function. It is interesting to note that (4.15) is not a member of Hayman's class of admissible functions. Actually (4.15) violates one of Hayman's assumptions that is not vital to the success of the method. This example provides an illustration of how Hayman's class of admissible functions could be extended along the lines of our present paper. The second relevant path contributes only to the second term of the asymptotic expansion.

5. Further preliminary results. From assumption 3(c) it is easily seen that $\log G(Z_k \exp u)$ is a regular function of the complex variable u providing $|u| < h\,(\log M_1)/n$. In this neighbourhood we may expand $\log G(Z_k \exp u)$ into a Maclaurin expansion and obtain

$$\log G(Z_k \exp u) = \log G(Z_k) + H \log G(Z_k)u + H^2 \log G(Z_k)\frac{u^2}{2!} + \sum_{m=3}^{\infty} H^m \log G(Z_k)\frac{u^m}{m!}. \tag{5.1}$$

Since $H \log G(Z_k) = n$ we may, by choosing $u = \log(z/Z_k)$, say

$$\log G(z) = \log G(Z_k) + n \log(z/Z_k) + H^2 \log G(Z_k)\frac{(\log z/Z_k)^2}{2!} + \sum_{m=3}^{\infty} H^m \log G(Z_k)(\log(z/Z_k))^m/m!, \tag{5.2}$$

providing $|\log(z/Z_k)| < h\,(\log M_1)/n$. If we denote the set of points satisfying $|\log(z/Z_k)| < h \log(M_1)/n$ by D_k, then the point $z = z_k(r)$ lies in D_k.

Assumption (3) tells us that

$$\lim_{r\to b} (\log G(Z_k \exp(\log M_1(r)\, w/n)) - \log G(Z_k))/\log M_1(r) = g_k(w), \tag{5.3}$$

and that the limit exists uniformly in w, for $|w| < h$. We may, therefore,

differentiate (5.3) m times with respect to w and then place $w = 0$. This gives

$$\lim_{r\to b} ((\log M_1(r))^{m-1} H^m \log G(Z_k)/n^m) = g_k^{(m)}(0). \tag{5.4}$$

From (5.4) the following results can easily be obtained.

$$g_k(0) = 0, \qquad g_k'(0) = 1. \tag{5.5}$$

$$R1\, g_k''(0) > 0. \tag{5.6}$$

There exist constants P_3 and P_4 such that

$$P_3 (\log M_1(r)) \leqslant |H^2 \log G(Z_k)| \leqslant P_4 ((\log M_1(r))^{1+2p}). \tag{5.7}$$

$$\lim_{r\to b} |H^2 \log G(Z_k)| = \infty. \tag{5.8}$$

$$\lim_{r\to b} \arg(H^2 \log G(Z_k)) = \arg g_k''(0). \tag{5.9}$$

$$H^2 \log G(Z_k) = 0(H^2 \log G(Z_1)). \tag{5.10}$$

$$|H^m \log G(Z_k)/m!| = 0(n^m/h^m (\log M_1)^{m-1}). \tag{5.11}$$

If we define ψ_k by

$$\psi_k = \lim_{r\to b} \arg(H^2 \log G(Z_k)) \tag{5.12}$$

then (5.6) and (5.9) tell us that

$$|\psi_k| < \tfrac{1}{2}\pi. \tag{5.13}$$

We shall define an infinitesimal ϵ by

$$\epsilon = |H^2 \log G(Z_1)|^{-\alpha} \tag{5.14}$$

where, for reasons that shall appear later, we restrict α to satisfy

$$(6p + 2)/6(2p + 1) < \alpha < \tfrac{1}{2}. \tag{5.15}$$

The constant p is taken from assumption (3). Since the $Rl\ (H^2 \log G(Z_k)) > 0$ we may introduce $(H^2 \log G(Z_k))^{\frac{1}{2}}$ in an unambiguous manner by taking the branch with a positive real part. The quantity $\phi_{0,k}(r)$ defined by

$$\phi_{0,k}(r) = |H^2 \log G(Z_1)|^{-\alpha} (H^2 \log G(Z_k))^{\frac{1}{2}} \tag{5.16}$$

has the properties that

$$|\phi_{0,k}(r)| \sim \frac{|g_k''(0)|^{\frac{1}{2}}}{|g_1''(0)|^{\alpha}} (n^2/\log M_1)^{\frac{1}{2}-\alpha} = 0[(\log M_1)^{(1+2p)(\frac{1}{2}-\alpha)}] \tag{5.17}$$

and

$$\lim_{r\to b} |\phi_{0,k}(r)| = \infty, \tag{5.18}$$

and

$$\lim_{r\to b} |\arg \phi_{0,k}(r)| = \tfrac{1}{2}|\psi_k| < \tfrac{1}{4}\pi. \tag{5.19}$$

Since

$$\lim_{r\to b} n\epsilon = \infty,$$

we see that $1/n$ is a higher order infinitesimal than ϵ. Further, the fact that

$$\lim_{r\to b} n\epsilon/\log M_1(r) = 0$$

implies that any point z lying on the line joining $z_k \exp i\epsilon$ to $Z_k \exp i\epsilon$ is a member of the domain D_k. For such points z we may prove

$$\log (z/Z_k) = i\epsilon + 0(1/n). \tag{5.20}$$

Further, since $1/n$ is a higher order infinitesimal than ϵ we may, to this order of approximation, write

$$\log (z/Z_k) = i\epsilon. \tag{5.21}$$

This fact plus (5.2) enables us to prove that

$$\log(G(z)/z^n) = \log(G(Z_k)/Z_k^n) - \frac{\epsilon^2}{2} H^2 \log G(Z_k) + 0(n^3\epsilon^3/(\log M_1)^2). \tag{5.22}$$

From (5.4)

$$\lim_{r\to b} [(\log M_1) H^2 \log G(Z_k)/n^2] = g_k''(0) \tag{5.23}$$

and from (5.6) $g_k''(0) \neq 0$. Hence

$$H^2 \log G(Z_k) = n^2(g_k''(0) + o(1))/\log M_1. \tag{5.24}$$

Thus (5.22) may be written

$$\log(G(z)\, Z_k^n/z^n\, G(Z_k)) = \frac{\epsilon^2}{2} H^2 \log G(Z_k)(1 + O(n\epsilon/\log M_1)). \tag{5.25}$$

We have already seen that

$$\lim_{r\to b} (n\epsilon/\log M_1) = 0.$$

Hence (5.25) can be shown to imply that a constant $P_5 > 0$ exists such that

$$|G(z)\, Z_k^n/G(Z_k)\, z^n| = O(\exp(-P_5(\log M_1)^{1-2\alpha})). \tag{5.26}$$

Hence $|G(z)\, Z_k^n/G(Z_k)z^n| \to 0$ with respect to $\log M_1(r)$. Since $G(Z_k)/M_1(r)$ and $(r/Z_k)^n$ are bounded functions of r this means

$$|G(z)\, r^n/M_1(r)z^n| \to 0, \text{ with respect to } \log M_1(r). \tag{5.27}$$

In particular (5.27) is true for the points $z_k \exp(i\epsilon)$ and $Z_k \exp(i\epsilon)$. Obviously the same result may be obtained by replacing ϵ by $-\epsilon$. We use (5.27) to prove the following lemma.

LEMMA 5.1. *The quantity $I(r)$ defined by*

$$I(r) = (r^n/M_1(r)) \int_C G(z)\, z^{-(n+1)}\, dz \tag{5.28}$$

is $= 0$, *with respect to* $\log M_1(r)$, *providing* L *is taken to be one of the following paths of integration.*

(a) The arc of the circle $|z| = r$ joining the points $z_k \exp(i\epsilon)$ to $z_{k+1} \exp(-i\epsilon)$
(b) The line joining $z_k \exp(i\epsilon)$ to $Z_k \exp(i\epsilon)$.
(c) The line joining $z_k \exp(-i\epsilon)$ to $Z_k \exp(-i\epsilon)$.

Proof. Our notation implies that $z_k(r)$, $z_{k+1}(r)$ are two points of intersection of two consecutive relevant paths with $z = r$. By assumption (2)

$$\lim_{r\to b} \arg z_k(r)$$

exist for $k = 1, 2, \ldots, N$ and, further, the values of these limits are distinct. Hence our numbering system can be assumed, without any loss of generality, to imply that

$$\lim_{r\to b} \arg z_k(r) < \lim_{r\to b} \arg z_{k+1}.$$

Further, since ϵ was defined by (5.14) we have

$$\lim_{r\to b} \arg(z_k(r) \exp(i\epsilon)) = \lim_{r\to b} \arg z_k(r).$$

This of course means that for r close to b

$$\arg z_k(r) < \arg (z_k(r) \exp (i\epsilon)) < \arg (z_{k+1}(r) \exp (-i\epsilon)) < \arg z_{k+1}(r).$$

This means that there can be no points of intersection of a relevant path and $|z| = r$ contained in $\arg (z_k(r) \exp (i\epsilon) \leqslant \theta \leqslant \arg (z_{k+1}(r) \exp (-i\epsilon)$. Thus on the path of integration given by part (a) we must have $|G(z)|$ less than or equal to the value of $|G(z)|$ at a maximum which is on an irrelevant path, or $|G(z)|$ is less than or equal to the value of $|G(z)|$ at one of the end-points. In either case $I(r) = 0$. To make our proof consistent even if the points z_N and z_1 are involved we adopt the convention that $|z_{N+1}| = |z_1|$ but $\arg z_{N+1} = \arg z_1 + 2\pi$.

For the straight line paths of integration (5.27) is sufficient to establish the result. As one would suspect, Lemma 5.1 enables us to recognize the portions of the contour C of (1.2) that can be discarded.

To complete our preliminary results we now proceed to study the behaviour of certain functions which enter into the derivation of the central theorem of the next section of our paper. Let ϕ be a complex variable and suppose we restrict ϕ to be on the line, in the complex ϕ plane, joining the points $\phi_{0,k}(r)$ to $-\phi_{0,k}(r)$ where $\phi_{0,k}$ is given by (5.16). We define $a_{m,k}$, $(m = 3, 4, \ldots, k = 1, 2, \ldots, N)$, by

$$(5.29) \qquad a_{m,k} = (\log M_1(r))^{\frac{1}{2}(m-2)} \mathrm{H}^m \log G(Z_k)\ (i\phi)^m / (m!(\tfrac{1}{2}\mathrm{H}^2 \log G(Z_k))^{m/2})$$

then (5.11) and (5.24) combined with (5.29) give

$$(5.30) \qquad |a_{m,k}| = 0((Q|\phi|)^m),$$

where Q is a positive constant. For this reason the functions $f_k(r, \phi, u)$, of the complex variable u, defined by

$$f_k(r, \phi, u) = \sum_{m=1}^{\infty} a_{m+2,k} u^m \tag{5.31}$$

are regular functions of u as long as $Q|\phi||u| < 1$. From the fact that $|\phi| < |\phi_{0,k}|$, (5.15) implies that if $|u| < 2(\log M_1(r))^{-\frac{1}{2}}$ then

$$\lim_{r \to b} |\phi u| = 0.$$

Thus the point $u = (\log M_1(r))^{-\frac{1}{2}}$ is certainly within the domain of regularity of $f_k(r, \phi, u)$. Let us now go on to consider the functions $F_k(r, \phi, u)$ defined by

$$F_k(r, \phi, u) = \exp f_k(r, \phi, u). \tag{5.32}$$

All of the functions $F_k(r, \phi, u)$ have Maclaurin expansions about $u = 0$ and the radii of convergence of these expansions are certainly all $> 2(\log M_1(r))^{-\frac{1}{2}}$ We write

$$F_k(r, \phi, u) = 1 + \sum_{m=1}^{\infty} b_{m,k}(r, \phi) u^m. \tag{5.33}$$

We can easily establish that

(5.34) $b_{2m+1,k}$ is a polynomial in ϕ and is an odd function of ϕ.

(5.35) $b_{2m,k}$ is a polynomial in ϕ and is an even function of ϕ.

(5.36) $|b_{m,k}| = 0((P_6|\phi|)^{2m})$, for large $|\phi|$, and P_6 and the order relation involve constants that are independent of r.

The result (5.36) follows from a lemma of Moser and Wyman **(7)**. Further, (5.15) implies that

$$\lim_{r \to b} |\phi_0|^3 (\log M_1(r))^{-\frac{1}{2}} = 0.$$

Hence as long as $|u| < 2(\log M_1(r))^{-\frac{1}{2}}$ we must have

$$\left| \sum_{s+1}^{\infty} b_{m,k}(r, \phi) u^m \right| = O((p_6|\phi|)^{3s+3} (\log M_1(r))^{-\frac{1}{2}(s+1)}). \tag{5.37}$$

From the fact that

$$\int_{-\infty \exp i\beta}^{\infty \exp i\beta} (\exp(-\phi^2)) \phi^m \, d\phi$$

exists as long as $|\beta| < \pi/4$ it is easily shown that

$$\left| \int_{-\phi_{0,k}}^{\phi_{0,k}} \left(\sum_{s+1}^{\infty} b_{m,k}(r, \phi) u^m \right) \exp(-\phi^2) \, d\phi \right| = O((\log M_1(r))^{-\frac{1}{2}(s+1)}) \tag{5.38}$$

where in (5.38) the constants entering into the order relation may depend on s but not on r. In particular if $R_{s,k}$ is given by

$$R_{s,k} = \int_{-\phi_{0,k}}^{\phi_{0,k}} \sum_{s+1}^{\infty} (b_{m,k}(r, \phi)\ (\log M_1(r))^{-\frac{1}{2}m})\ \exp(-\phi^2)\, d\phi \tag{5.39}$$

then

$$|R_{s,k}| = 0((\log M_1(r))^{-(s+1)}). \tag{5.40}$$

This completes the preliminary results necessary for the proof of our central theorem.

6. The asymptotic formula. From (1.2) the Laurent coefficients a_n are given by

$$a_n = (2\pi i)^{-1} \int_C G(z)\, z^{-(n+1)}\, dz, \tag{6.1}$$

where C can be any contour within the domain of regularity of $G(z)$ and enclosing $z = 0$. We shall assume n is a large positive integer. Hence the convention of § 4 that a choice of n determines the value $r = r(n) = |Z_1(n)|$, where $Z_1(n)$ is the featured solution of

$$H \log G(z) = n. \tag{6.2}$$

We may also consider the points $z_k(r)$, $Z_k(n)$, $(k = 1, 2, \ldots, N)$ as determined. The contour C is now chosen as follows.

Contour C. The points in the z plane $Z_k \exp(-i\epsilon)$, $Z_k \exp(i\epsilon)$ are joined by arcs of the circles $|z| = |Z_k|$. The points $Z_k \exp(-i\epsilon)$ and $Z_k \exp(i\epsilon)$ are joined, respectively, to the points $z_k \exp(-i\epsilon)$ and $z_k \exp(i\epsilon)$ by straight lines. Finally $z_k \exp(i\epsilon)$ is joined to $z_{k+1} \exp(-i\epsilon)$ by the arc of the circle $|z| = |z_k| = r$. The convention adopted in a previous section concerning z_{N+1} assures that C becomes a closed contour. Further, the previous work on relevant paths also ensures that the contour C does not traverse any of its points more than once.

From Lemma 5.1 the only portions of C that can contribute to the asymptotic expansion of $r^n a_n / M_1(r)$ are those portions that intersect a relevant path. For this reason we may write

$$r^n a_n / M_1(r) \sim \sum_{k=1}^{N} I_k \tag{6.3}$$

where

$$I_k = (2\pi i M_1(r))^{-1} r^n \int_{C_k} G(z)\, z^{-(n+1)}\, dz \tag{6.4}$$

and C_k is the arc of the circle $|z| = |Z_k|$ joining $Z_k \exp(-i\epsilon)$ to $Z_k \exp(i\epsilon)$. For this reason we study

$$K_k = (2\pi i)^{-1} \int_{C_k} G(z)\, z^{-(n+1)}\, dz. \tag{6.5}$$

The substitution $z = Z_k \exp i\theta$ makes (6.5) become

$$K_k = (2\pi)^{-1} Z_k^{-n} \int_{-\epsilon}^{\epsilon} G(Z_k \exp i\theta) \exp(-in\theta)\, d\theta. \tag{6.6}$$

Expanding $\log G(Z_k \exp i\theta)$ in a Maclaurin expansion about $\theta = 0$ we have

$$\log G(Z_k \exp i\theta) - in\theta = \log G(Z_k) - \tfrac{1}{2} H^2 \log G(Z_k)\, \theta^2 + \sum_{m=3}^{\infty} H^m \log G(Z_k)(i\theta)^m/m! \tag{6.7}$$

Hence the substitution

$$\phi = (\tfrac{1}{2} H^2 \log G(Z_k))^{\frac{1}{2}} \theta \tag{6.8}$$

into (6.6) yields

$$2\pi K_k = G(Z_k) Z_k^{-n} (\tfrac{1}{2} H^2 \log G(Z_k))^{-\frac{1}{2}} \int_{-\phi_{0,k}}^{\phi_{0,k}} F_k(r, \phi, (\log M_1(r))^{-\frac{1}{2}}) \exp(-\phi^2)\, d\phi \tag{6.9}$$

where $\phi_{0,k}$ is given by (5.16) and $F_k(r, \phi, u)$ by (5.32). From (5.33) and (5.39) we may write

$$\begin{aligned} &\int_{-\phi_{0,k}}^{\phi_{0,k}} F_k(r, \phi, (\log M_1(r))^{-\frac{1}{2}}) \exp(-\phi^2)\, d\phi \\ &= \int_{-\phi_{0,k}}^{\phi_{0,k}} \exp(-\phi^2)(1 + \sum_{m=1}^{2s} b_{m,k}(r, \phi)(\log M_1(r))^{-m/2})\, d\phi + R_{2s,k} \\ &= \int_{-\phi_{0,k}}^{\phi_{0,k}} \exp(-\phi^2)(1 + \sum_{m=1}^{s} b_{2m,k}(\log M_1(r))^{-m})\, d\phi + O((\log M_1(r))^{-s-\frac{1}{2}}) \end{aligned} \tag{6.10}$$

because of (5.34) and (5.40). Further, since $b_{2m,k}(r, \phi)$ is a polynomial in ϕ we include only terms which are exponentially small by replacing $\phi_{0,k}$ by $\infty \exp(i\psi_k)$, and without any loss in generality we may consider ϕ to be now a real variable and take limits of integration from $-\infty$ to ∞. If we let $A_{m,k}(r)$ be given by

$$A_{m,k}(r) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp(-\phi^2)\, b_{2m,k}(r, \phi)\, d\phi \tag{6.11}$$

then we have proven

$$\int_{-\phi_{0,k}}^{\phi_{0,k}} F_k(r, \phi, (\log M_1(r))^{-\frac{1}{2}}) \exp(-\phi^2)\, d\phi \sim \sqrt{\pi}\Big(1 + \sum_{m=1}^{\infty} A_{m,k}(\log M_1(r))^{-m}\Big) \tag{6.12}$$

and (6.9) gives

$$K_k \sim G(Z_k) Z_k^{-n} (2\pi H^2 \log G(Z_k))^{-\frac{1}{2}} \Big(1 + \sum_{m=1}^{\infty} A_{m,k}(\log M_1(r))^{-m}\Big). \tag{6.13}$$

This in turn gives the complete asymptotic expansion for I_k in (6.4) and a_n from (6.3). Although by assumption $G(Z_k)/M_1(r)$ is not asymptotically equal to zero, it is possible that

$$G(Z_k)/M_1(r) = O\left(\frac{1}{(\log M_1(r))^s}\right),$$

and in such a case I_1 and I_k would be terms of different orders. For this reason it is better to leave (6.13) as our major formula and in each specific case write out (6.3) after the different orders of the K_k have been established. We shall illustrate the procedure in the next section. Thus we have established the following central theorem of our paper.

CENTRAL THEOREM. *A complete asymptotic expansion of the Laurent coefficients of any generating function $G(z)$ that is admissible can be obtained by the formulae established in this section.*

In fact it is possible to show that the first two terms of (6.13) are given by

$$K_k \sim \frac{G(Z_k)}{Z_k^n(2\pi H^2 \log G(Z_k))^{\frac{1}{2}}} \left[1 + \frac{3H^2 \log G(Z_k)\, H^4 \log G(Z_k) - 5(H^3 \log G(Z_k))^2}{24\,(H^2 \log G(Z_k))^3}\right]. \tag{6.14}$$

In using these formulae we must examine the solutions of

$$H \log G(z) = n \tag{6.15}$$

and from these we must select the featured path $z = Z_1(n)$ and establish all of the relevant paths $z = Z_k(n)$, $(k = 1, 2, \ldots, N)$. The variable r is determined by

$$r = |Z_1(n)| \tag{6.16}$$

and all quantities involved in the asymptotic expansion may then be computed.

7. Specific examples. Since we have already shown that $[\exp(\frac{1}{2}(1 - z^2))]/(1 - z)$ is an admissible function we shall use this generating function as our first illustration. From (4.20) and (4.21) we have

$$r = Z_1(n) = 1 - \frac{1}{2n^{\frac{1}{2}}} - \frac{3}{8n} + \frac{1}{8n^{3/2}} + \ldots \tag{7.1}$$

$$Z_2(n) = -1 + \frac{1}{2n^{\frac{1}{2}}} - \frac{1}{8n} - \frac{1}{8n^{3/2}} + \ldots \tag{7.2}$$

Thus

$$M_1(r) = 2n^{\frac{1}{2}}\, e^{(2n^{\frac{1}{2}}-1)/4}\left(1 - \frac{5}{16n^{\frac{1}{2}}} + \ldots\right) = G(Z_1). \tag{7.3}$$

$$G(Z_2) = \tfrac{1}{2}e^{(2n^{\frac{1}{2}}+1)/4}\left(1 + \frac{7}{16n^{\frac{1}{2}}} + \ldots\right). \tag{7.4}$$

$$H^2\log G(Z_1) = 4n^{3/2}\left(1 - \frac{2}{n^{\frac{1}{2}}}\right). \tag{7.5}$$

$$H^2\log G(Z_2) = 4n^{3/2}\,(1 + \ldots). \tag{7.6}$$

$$H^3\log G(Z_1) = 24n^2\,(1 + \ldots). \tag{7.7}$$

$$H^4\log G(Z_1) = 192\,n^{5/2}\,(1 + \ldots). \tag{7.8}$$

$$K_1 \sim \frac{e^{n^{\frac{1}{2}}+1/4}}{(2\pi)^{\frac{1}{2}}n^{1/4}}\left(1 + \frac{5}{12n^{\frac{1}{2}}} + \ldots\right). \tag{7.9}$$

$$K_2 \sim \frac{(-1)^n\, e^{n^{\frac{1}{2}}+0.25}}{4(2\pi)^{\frac{1}{2}}n^{3/4}}\,(1 + \ldots). \tag{7.10}$$

Thus we have an illustration of a case where K_2 is a different order term than K_1 and K_2 affects only the second term of the asymptotic expansion. From (7.9) and (7.10) we find

$$a_n \sim \frac{e^{n^{\frac{1}{2}}+1/4}}{(2\pi)^{\frac{1}{2}}n^{1/4}}\left(1 + \left(\frac{5}{12} + \frac{(-1)^n}{4}\right)\frac{1}{n^{\frac{1}{2}}} + \ldots\right). \tag{7.11}$$

We have so far in the paper only used admissible functions which have Maclaurin expansions. One of the simplest examples of an admissible function that has no Maclaurin expansion about $z = 0$ and does have a Laurent expansion is the generating function $\exp\,(\frac{1}{2}x(z - z^{-1}))$ for the Bessel Functions $J_n(x)$. We shall assume x is fixed, real, and positive. To use our method on this particular generating function is a little like shooting sparrows with cannons. It is of course well known that the series definition of $J_n(x)$ will act as an asymptotic expansion under these conditions. However it does act to illustrate the procedure for examples of similar type but more complicated in nature. For this generating function there is only one relevant path $\theta = 0$.

$$r = |Z_1(n)| = \frac{n + (n^2 - x^2)^{\frac{1}{2}}}{x}. \tag{7.12}$$

Hence

$$M_1(r) = G(Z_1) = \exp\,[(n^2 - x^2)^{\frac{1}{2}}]. \tag{7.13}$$

$$H^2\log G(Z_1) = (n^2 - x^2)^{\frac{1}{2}}. \tag{7.14}$$

$$H^3\log G(Z_1) = n. \tag{7.15}$$

$$H^4\log G(Z_1) = (n^2 - x^2)^{\frac{1}{2}}. \tag{7.16}$$

$$J_n(x) \sim \frac{x^n \exp[(n^2 - x^2)^{\frac{1}{2}}]}{2^{\frac{1}{2}}\pi^{\frac{1}{2}}(n^2 - x^2)^{1/4}[n + (n^2 - x^2)^{\frac{1}{2}}]^n}\left(1 - \frac{2n^2 + 3x^2}{24(n^2 - x^2)^{3/2}} + \ldots\right). \tag{7.17}$$

As a final example we use the generating function $\exp(z/(1+z^2))$ to illustrate the method when the relevant paths are not straight lines. For this generating function there are two relevant paths. These are

$$Z_1(n) = i + \frac{1-i}{2n^{\frac{1}{2}}} - \frac{1}{4n} + \dots . \tag{7.18}$$

$$Z_2(n) = -i + \frac{1+i}{2n^{\frac{1}{2}}} - \frac{1}{4n} + \dots = \overline{Z}_1(n). \tag{7.19}$$

Since $z_1(r) = Z_1(n)$, $z_2(r) = Z_2(n)$ the fact that $\exp(z/(1+z^2))$ is admissible can easily be verified. Hence

$$G(Z_1) = [\exp(\tfrac{1}{2}(n^{\frac{1}{2}}(1+i)))](1+\dots). \tag{7.20}$$

$$G(Z_2) = [\exp(\tfrac{1}{2}(n^{\frac{1}{2}}(1-i)))](1+\dots). \tag{7.21}$$

$$Z_1^{-s} = [\exp(\tfrac{1}{2}(n^{\frac{1}{2}}(1+i) - i\,n\pi))](1+\dots). \tag{7.22}$$

$$Z_2^{-s} = [\exp(\tfrac{1}{2}(n^{\frac{1}{2}}(1-i) + in\pi))](1+\dots). \tag{7.23}$$

$$H^2 \log G(Z_1) = 2(1-i)n^{3/2}(1+\dots). \tag{7.24}$$

$$H^2 \log G(Z_2) = (2(1+i)n^{3/2}(1+\dots). \tag{7.25}$$

and

$$a_n \sim \frac{2e^{n^{\frac{1}{2}}}\cos[n^{\frac{1}{2}} - \frac{1}{8}(4n-1)\pi]}{(2\pi)^{\frac{1}{2}}(2n)^{3/4}}. \tag{7.26}$$

8. Conclusion. The major result of this paper has been to give a set of conditions on a complex function $G(z)$ by means of which we can recognize whether or not $G(z)$ belongs to our class of admissible functions. If $G(z)$ does belong to such a class then our central theorem tells us that the complete asymptotic behaviour of the Laurent coefficients can be determined. It would be desirable to have results that tell us that certain large classes of functions are admissible. For example, it is possible to show that all functions of the form $P_m(z)\exp(S_t(z) + Q(1/z))$ are admissible providing $P_m(z)$, $S_k(z)$ are polynomials and $Q(1/z)$ is a regular function of its argument. From our experience we would say that there exists a very extensive class of integral functions all of which are admissible. Problems of this type would be worth investigation so that in specific examples one could tell, almost at a glance, whether or not the generating function is admissible. In such cases one can almost use our central theorem to write down the asymptotic behaviour of the Laurent coefficients.

In our terminology Hayman's class of admissible functions required that $G(z)$ possess a unique relevant path $\theta = 0$ that contributes to the first term of the asymptotic expansion. For this reason he was able to prove results stating that the product of admissible functions is admissible. In our case the two functions involved might have quite different relvant paths and

such a theorem is not possible. For example, the functions $\exp(z/(1-z^2))$ and $(\exp(z/(z^2-1)))/(1-z)$ are both admissible but their product $1/(1-z)$ is not admissible.

When one considers the class of functions for which the method of Darboux applies and the class of functions for which the method outlined in this paper applies, we now have a very extensive class of functions for which we can consider the problem of determining the behaviour of a_n, as $n \to \infty$, as a solved problem.

References

1. G. M. Darboux, *Mémoire sur l'approximation des fonctions de très-grandes nombres, et sur une classe étendue de développements en série*, J. Math. pur. et appl., *4* (1878), 6–56; 377–416.
2. A. Erdélyi, *Asymptotic expansions* (New York, 1956).
3. G. H. Hardy, *The maximum modulus of an integral function*, Quart. J. Math., *XLI* (1909), 1–9.
4. W. K. Hayman, *A generalization of Stirling's formula*, J. reine und angew., *196* (1956), 67–95.
5. L. Moser and M. Wyman, *Asymptotic expansions*, Can. J. Math., *8* (1956), 225–33.
6. ———, *Asymptotic expansions, II*, Can. J. Math., *9* (1957), 194–209.
7. ———, *On solutions of $x^d = 1$ in symmetric groups*, Can. J. Math., *7* (1955), 159–68.
8. G. Szekeres, *Some asymptotic formulae in the theory of partitions, II*, Quart. J. Math., Oxford, Ser. (2), *4* (1953), 97–111.
9. G. Valiron, *Lectures on the general theory of integral functions* (New York, 1949).

University of Alberta

CONVEX STRUCTURES AND CONTINUOUS SELECTIONS

ERNEST MICHAEL

Introduction. This paper continues the study of continuous selections begun in **(13; 14; 15)** and the expository paper **(12)**.[1] The purpose of these papers, which is described in detail in the introduction to **(13)**, can be summarized here as follows. If X and Y are topological spaces, and ϕ a function (called a *carrier*) from X to the space 2^Y of non-empty subsets of Y, then a *selection* for ϕ is a continuous $f: X \to Y$ such that $f(x) \in \phi(x)$ for every $x \in X$. For reasons which are explained in **(13)**, we restrict our attention to carriers which are *lower semi-continuous* (l.s.c.), in the sense that, whenever U is open in Y, then $\{x \in X \mid \phi(x) \cap U \neq \emptyset\}$ is open in X. Our purpose in these papers is to find conditions for the existence and extendability of selections.

The principal purpose of *this* paper is to generalize the following result, which is half of the principal theorem (Theorem 3.2″) of **(13)** (and is repeated as Theorem I of **(12)**).

THEOREM A. *Let S be paracompact, Y a Banach space, and $\mathfrak{S}$ the family of closed, convex, non-empty subsets of Y. Then every lower semi-continuous carrier $\phi: X \to \mathfrak{S}$ admits a selection.*

In this paper, the Banach space Y is replaced by a complete metric space carrying an axiomatically defined *convex structure* (Definition 1.1) which permits one to take "convex combinations" of some (but not necessarily all) ordered n-tuples of points in Y, in a suitably continuous fashion. With *convex* sets defined in the obvious way (Definition 1.4), the generalization of Theorem A which is thus obtained is given in Theorem 1.5 (a). The significance of this generalization is illustrated by the following example.

Example A. G is a metrizable group, and H a closed subgroup which is isomorphic to the additive group of a Banach space. Convex combinations can be taken of n-tuples lying in the same right coset of H. The right cosets of H are then convex sets. By applying Theorem 1.5 (a) to this situation, it is shown in Corollary 7.3 that there exists a cross-section.[2] (For G a Banach

Received May 20, 1958. This paper was written in part on a National Science Foundation contract at the Institute for Advanced Study, and in part on a National Science Foundation contract at the University of Washington.

[1] Except for § 8, no previous knowledge of continuous selections is necessary to read this paper, although some acquaintance with the first three sections of **(13)** will be helpful.

[2] If G is a topological group, H a closed subgroup, u the canonical map from G to G/H, and if $\phi: G/H \to G$ is defined by $\phi(x) = u^{-1}(x)$, then a *cross-section* (resp. *local cross section*) is a selection for ϕ (resp. $\phi|U$ for some non-empty open $U \subset G/H$).

space and H a linear subspace, this already follows from Theorem A, and was first proved by Bartle and Graves in **(2)**.)

If the family of sets $\mathfrak{S}$, instead of having convex elements, is only what we shall call (Definition 1.4) *equi-locally convex*, then the global Theorem 1.5 (a) changes into the local Theorem 1.5 (b). This is illustrated by the following example.

Example B. If Example A is modified only by assuming H to be *locally* isomorphic to the additive group of a Banach space, then only "close together" n-tuples in the same right coset of H admit convex combinations, the collection of right cosets of H is equi-locally convex, and Theorem 1.5 (b) implies that there exists a *local* cross-section[3] (Corollary 7.3).

Both parts of Theorem 1.5 will be obtained as simple consequences of a single Theorem, 1.3, which also generalizes Theorem A. The statements of these theorems, as well as the basic definitions, are found in § 1. § 2 contains some preliminary results which are not directly concerned with selections. § 3, the core of this paper, contains the proof of Theorem 1.3. This proof parallels that of Theorem A in **(13)**, by obtaining the desired function f as the limit of a uniformly convergent sequence of functions f_n, However, while in the proof of Theorem A the f_n's are *continuous* with $f_n(x)$ *near* $\phi(x)$ for all x, in the proof of Theorem 1.3 the f_n's may be *discontinuous* with $f_n(x)$ *in* $\phi(x)$ for all x; this new approach is necessitated by the inability to take convex combinations of n-tuples in Y not lying in an element of $\mathfrak{S}$. The simple derivation of Theorem 1.5 from Theorem 1.3 is contained in § 4.

An important source of convex structures on a metric space Y is provided by so-called *geodesic structures* on Y, which permit one to take convex combinations of certain *pairs* of points of Y in an appropriate fashion. Geodesic structures are studied in § 5, where it is shown (Proposition 5.3) how they inductively generate convex structures in a canonical fashion. In § 6, a recent theorem of Nijenhuis **(17)** is used to show that the usual geodesic segments on a Riemannian manifold give rise to a geodesic structure, and hence to a convex one.

The last two sections are devoted to applications of Theorems 1.3 and 1.5 and, except when dealing with Lie groups, are independent of §§ 5 and 6. § 7 deals with *locally convex groups*, which are, essentially, topological groups with an invariant convex structure defined on a neighbourhood of the identity. Theorem 1.5 is used to prove a cross-section theorem (Theorem 7.2) for such groups, which implies (Corollary 7.3) the results given in Examples A and B above, as well as Gleason's cross-section theorem for arbitrary Lie groups. § 8, finally, uses Theorem 1.3 to prove a theorem of the covering homotopy type, which is then applied to fibre spaces in the sense of Hurewicz **(7)**.

In conclusion, it should be noted that the results of this paper are of interest only if the domain X of the functions to be defined is *infinite* dimen-

[3] *Ibid.*

sional; for *finite* dimensional X, better results were already obtained in **(14)** and, with reference to § 8, in **(15)**.

1. Principal theorems and definitions. Throughout this paper, P_n denotes the unit simplex in Euclidean n-space R^n; that is,

$$P_n = \left\{ t \in R^n \mid 0 \leqslant t_i \leqslant 1, \quad (i = 1, \ldots, n), \sum_{i=1}^{n} t_i = 1 \right\}.$$

If E is any set, then E^n will denote the n-fold Cartesian product of E, and if $i \leqslant n$, then $\partial_i : E^n \to E^{n+1}$ is defined by $\partial_i(x_1, \ldots, x_n) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

Definition 1.1. A *convex structure* on a metric space E with metric ρ assigns to each positive integer n a subset M_n of E^n, and a function $k_n : M_n \times P_n \to E$, such that

(a) If $x \in M_1$, then $k_1(x, 1) = x$.

(b) If $x \in M_n$ $(n > 2)$ and $i \leqslant n$, then $\partial_i x \in M_{n-1}$ and, for any $t \in P_n$ with $t_i = 0$, $k_n(x, t) = k_{n-1}(\partial_i x, \partial_i t)$.

(c) If $x \in M_n$ $(n > 2)$ with $x_i = x_{i+1}$ for some $i < n$, and if $t \in P_n$, then $k_n(x, t) = k_{n-1}(\partial_i x, t^*)$, where $t^* = (t_1, \ldots, t_{i-1}, t_i + t_{i+1}, t_{i+2}, \ldots, t_n)$.

(d) If $x \in M_n$, then the map $t \to k_n(x, t)$, from P_n to E, is continuous.

(e) For all $\epsilon > 0$ there exists a neighbourhood V_ϵ of the diagonal in $E \times E$ such that, for all n and all $x, y \in M_n$, $(x_i, y_i) \in V_\epsilon$ for $i = 1, \ldots, n$ implies $\rho(k_n(x, t), k_n(y, t)) < \epsilon$ for all $t \in P_n$.

Note that conditions (a) and (c) together imply that, if $x \in M_n$ with $x_1 = \ldots = x_n$, then $k_n(x, t) = x_1$ for all $t \in P_n$.

Definition 1.2. A subset S of a space E with convex structure is *admissible* if $S^n \subset M_n$ for all n. If S is admissible, then the *convex hulk of* S, denoted by $\mathrm{conv}(S)$, is

$$\{k_n(x, t) \mid x \in S^n, \quad t \in P_n, \quad n = 1, 2, \ldots\}.$$

THEOREM 1.3. *Let Y be a complete metric space with a convex structure, and let $\mathfrak{S}$ be the family of non-empty admissible subsets of Y. Let X be paracompact, and $\phi : X \to \mathfrak{S}$ lower semi-continuous. Then there exists a continuous $f : X \to Y$ such that*

$$f(x) \in [\mathrm{conv}(\phi(x))]^-$$

for all $x \in X$.

Definition 1.4. Let E be a metric space with convex structure. Then a subset S of E is *convex* if it is admissible and $\mathrm{conv}(S) \subset S$. A family $\mathfrak{S}$ of subsets is *equi-locally convex* if there exists an open covering $\mathfrak{B}$ of E such that, whenever $S \in \mathfrak{S}$ and $B \in \mathfrak{B}$, then $S \cap B$ is admissible and $\mathrm{conv}(S \cap B) \subset S$.

Note that every family of convex sets is equi-locally convex.

THEOREM 1.5. *Let X be paracompact, Y a metric space with convex structure, $\mathfrak{S}$ a family of non-empty, complete subsets of Y, and $\phi: X \to \mathfrak{S}$ lower semi-continuous. Let $A \subset X$ be closed and let g be a selection for $\phi|A$. Then*

(a) *If every $S \in \mathfrak{S}$ is convex, then g can be extended to a selection for ϕ.*

(b) *If $\mathfrak{S}$ is equi-locally convex, then g can be extended to a selection for $\phi|U$ for some open $U \supset A$.*

No discussion of spaces with a convex structure would be complete without mentioning their extension properties. For convenience, let us call a metric space E *convex* if it admits a convex structure making E itself a convex set. As a special case of Theorem 1.5, we see that a convex complete metric space is an AE (absolute extensor) for paracompact spaces (in the sense of **(10)**); a somewhat better result is obtainable directly as a straightforward generalization of Dugundiji's extension theorem **(5**, Theorem 4.1**)**. A more interesting fact is a partial converse: *Every compact metric space E, which is an* AE *for metric spaces, is convex.* To see this, embed E in a Banach space B **(8)**, and let F be the closed convex hull of E in B. Then there exists a retraction $r: F \to E$, and we define $k_n: E^n \times P_n \to E$ by

$$k_n(x, t) = r\left(\sum_{i=1}^{n} t_i x_i\right).$$

That conditions (a) — (d) of Definition 1.1 are satisfied is clear, while (e) follows from the factthat F is compact **(9)** and hence r uniformly continuous. It is not know whether compactness can be replaced by a weaker condition in the above result.

2. Two lemmas. Our first lemma deals with the following very elementary concept.

Definition 2.1. If X is a topological space, (Y, ρ) a metric space, and $a \geq 0$, then a function $f: X \to Y$ is *a-continuous at $x_0 \in X$* if to every $\epsilon > 0$ there corresponds a neighbourhood U of x_0 such that

$$\rho(f(x), \ f(x_0)) < a + \epsilon$$

for every x in U. Moreover, f is *a-continuous* if it is a-continuous at every $x_0 \in X$.

Note that clearly f is 0-continuous (at x_0) if, and only if, f is continuous (at x_0).

LEMMA 2.2. *Le X be a topological space, (Y, ρ) a metric space, and let $f_n: X \to Y$ $(n = 1, 2, \ldots)$ be a sequence of functions which converges uniformly to a function $f: X \to Y$. Let $x_0 \in X$, and suppose that to each $a > 0$ and positive integer N corresponds an $n > N$ such that f_n is a-continuous at x_0. Then f is continuous at x_0.*

Proof. Let $\epsilon > 0$. Pick an n such that $\rho(f_n(x), f(x)) < \frac{1}{3}\epsilon$ for all $x \in X$, and

f_n is $\frac{1}{3}\epsilon$-continuous at x_0. Pick a neighbourhood U of x_0 such that, if $x \in U$, then $\rho(f_n(x), f_n(x_0)) < \frac{1}{3}\epsilon$. Then, if $x \in U$,

$$\rho(f(x_0), f(x)) \leqslant \rho(f(x_0), f_n(x_0)) + \rho(f_n(x_0), f_n(x) + \rho(f_n(x), f(x)) < \epsilon.$$

This completes the proof.

Our second lemma concerns partitions of unity.[4] In it, as well as later in the paper, we denote the set of points where a real-valued function f is positive by $P(f)$. An indexed family of real-valued functions $\{p_\alpha\}_{\alpha \in A}$ will be called *locally-finite* if the indexed family of sets $\{P(p_\alpha)\}_{\alpha \in A}$ is locally finite.[5]

LEMMA 2.3. *Let X be paracompact, Y a topological space, $\phi: X \to 2^Y$ lower semi-continuous, and $\{\sigma_n\}_{n=1}^{\infty}$ a sequence of continuous pseudometrics on Y. Then for each integer $n > 0$ there exists an index set A_n, a locally finite partition of unity $\{p_\alpha\}_{\alpha \in A_n}$ on X, points $y_\alpha(x) \in \phi(x)$ whenever $\alpha \in A_n$ and $x \in P(p_\alpha)$, and a map $\pi_n: A_{n+1} \to A_n$ onto, such that the following conditions are satisfied for all n.*

(a) *If $\alpha \in A_n$ and $x, x' \in P(p_\alpha)^-$, then $\sigma_n(y_\alpha(x), y_\alpha(x')) < 1$.*

(b) *If $\alpha \in A_n$, $\beta \in \pi_n^{-1}(\alpha)$, and $x \in P(p_\beta)^-$, then $\sigma_n(y_\alpha(x), y_\beta(x)) < 1$.*

(c) *If $\alpha \in A_n$, then $p_\alpha(x) = \Sigma\{p_\beta(x) | \beta \in \pi_n^{-1}(\alpha)\}$ for all $x \in X$.*

Proof. This theorem follows easily from **(16**, Lemma 2.1**)**. The statement of that result parallels that of our theorem, with the following differences: Instead of a partition of unity $\{p_\alpha\}_{\alpha \in A_n}$ on X, **(16**, Lemma 2.1**)** only provided us with an open covering $\{U_\alpha\}_{\alpha \in A_n}$ of X, the sets $P(p_\alpha)^-$ in (a) and (b) were replaced by U_α, and (c) was replaced by

(c′) *If $\alpha \in A_n$, then $U_\alpha = \cup\{U_\beta | \beta \in \pi_n^{-1}(\alpha)\}$.*

To obtain our partitions of unity, we proceed by induction, in such a way that

(d) $P(p_\alpha)^- \subset U_\alpha \qquad \alpha \in A_n,\ n = 1, 2, \ldots.$

For $n = 1$, simply pick an open covering $\{V_\alpha\}_{\alpha \in A_1}$ of X such that $\bar{V}_\alpha \subset U_\alpha$ for all $\alpha \in A_1$, and let $\{p_\alpha\}_{\alpha \in A_1}$ be a partition of unity subordinated[6] to $\{V_\alpha\}_{\alpha \in A_1}$. Suppose we have $\{p_\alpha\}_{\alpha \in A_n}$, and let us construct $\{p_\beta\}_{\beta \in A_{n+1}}$. For convenience, we shall write π for π_n.

Let $\alpha \in A_n$. Then $\{U_\beta \cap P(p_\alpha)^-\}_{\beta \in \pi^{-1}(\alpha)}$ is a relatively open covering of $P(p_\alpha)^-$ by (c′) and (d), and hence has a relatively open refinement $\{V_\beta\}_{\beta \in \pi^{-1}(\alpha)}$ with $\bar{V}_\beta \subset U_\beta$ for all $\beta \in \pi^{-1}(\alpha)$. Let $\{q_\beta\}_{\beta \in \pi^{-1}(\alpha)}$ be a partition of untiy on

---

[4] *A partition of unity* in a topological space is an indexed family $\{p_\alpha\}_{\alpha \in A}$ of functions from X to $[0, 1]$ such that $\Sigma_{\alpha \in A} p_\alpha(x) = 1$ for all $x \in X$.

[5] An indexed family of sets $\{U_\alpha\}_{\alpha \in A}$ is called *locally finite* if each $x \in X$ has a neighbourhood intersecting U_α for only finitely many $\alpha \in A$.

[6] $\{p_\alpha\}_{\alpha \in A}$ is *subordinated to* $\{U_\alpha\}_{\alpha \in A}$ if p_α vanishes outside U_α for all α. Every open covering of a paracompact space has a partition of unty subordinated to it.

$P(p_\alpha)^-$ which is subordinated to $\{V_\beta\}_{\beta\in\pi^{-1}(\alpha)}$. For each $\beta \in \pi^{-1}(\alpha)$, define $p_\beta: X \to [0, 1]$ by

$$p_\beta(x) = 0 \qquad x \notin P(p_\alpha),$$
$$p_\beta(x) = p_\alpha(x) \cdot q_\beta(x) \qquad x \in P(p_\alpha)^-.$$

This function p_β is well defined, and is continuous on X since it is continuous on two closed sets whose union is X. Moreover, $\{p_\beta\}_{\beta\in A_{n+1}}$ is now clearly a partition of unity on X which satisfies (a) – (d). This completes the proof.

3. Proof of Theorem 1.3. We begin by observing that conditions (c) and (e) of Definition 1.1 are equivalent to the following apparently stronger conditions (c′) and (e′). In fact, (c′) follows from (c) by straightforward induction, while (e′) follows from (e) by means of a standard result (see, for instance **(11**, Corollary 2.3**)**).

1.1 (c′). Let m and m' be positive integers with $m \leqslant m'$, $y \in M_m$, $t \in P_m$, $y' \in M_{m'}$, and $t' \in P_{m'}$. Suppose that, for some order-preserving map $\pi: \{1, \ldots, m'\} \to \{1, \ldots, m\}$, we have

$$y'_i = y_{\pi(i)} \qquad i = 1, \ldots, m',$$
$$t_j = \sum_{i\in\pi^{-1}} (j)t'_i \qquad j = 1, \ldots, m.$$

Then

$$k_m(y, t) = k_{m'}(y', t').$$

1.1 (e′). For all integers $n > 0$ there exsits a continuous pseudometric σ_n on E with the following property: If $x, y \in M_m$ for some m, then $\sigma_n(x_i, y_i) < 1$ for $i = 1, \ldots, m$ implies $\rho(k_m(x, t), k_m(y, t)) < 2^{-n}$ for all $t \in P_m$.

The main step in the roof of Theorem 1.3 will be the following lemma, whose proof, in turn, rests heavily on Lemma 2.3.

LEMMA 3.1. *With X, Y, and $\phi: X \to 2^Y$ as in Theorem* 1.3, *there exists a sequence of functions $f_n: X \to Y$ $(n = 1, 2, \ldots)$ such that, for all n,*

(a) *$f_n(x) \in \operatorname{conv}(\phi(x))$ for all $x \in X$,*
(b) *f_n is 2^{-n}-continuous,*
(c) *$\rho(f_n(x), f_{n+1}(x)) < 2^{-n}$ for all $x \in X$.*

Proof. We begin by applying Lemma 2.3, with σ_n $(n = 1, 2, \ldots)$ as in 1.1(e′) at the beginning of this section. Let A_n, p_α, $y_n(x)$, and π_n be as in Lemma 2.3. Let us also suppose that the sets A_n are well ordered in such a way that each π_n is order preserving; this is easily done by induction.

For each n and $x \in X$, let

$$A_n(x) = \{\alpha \in A_n \mid x \in P(p_\alpha)\},$$
$$\bar{A}_n(x) = \{\alpha \in A_n \mid x \in P(p_\alpha)^-\}.$$

Note that $A_n(x) \subset \bar{A}_n(x)$, and that both are finite. Let $\alpha_1, \ldots, \alpha_m$ be the

elements of $\bar{A}_n(x)$ in order. Remembering that $\phi(x)$ is admissible, we can now define $f_n(x)$ by

$$f_n(x) = k_m((y_{\alpha_1}(x), \ldots, y_{\alpha_m}(x)), (p_{\alpha_1}(x), \ldots, p_{\alpha_m}(x))). \tag{1}$$

Note that, since $p_\alpha(x) = 0$ if $\alpha \notin A_n(x)$, it follows from Definition 1.1(b) that $f_n(x)$ is not changed by the omission from (1) of terms with index $\alpha \notin A_n(x)$.

Let us check that our conditions are satisfied.

(a) This follows from the definitions.

(b) We will show that f_n is 2^{-n}-continuous at a given $x_0 \in X$. Let

$$V = \cap\{P(p_\alpha) | \alpha \in A_n(x_0)\} - \cup\{P(p_\alpha)^- | \alpha \in (A_n - \bar{A}_n(x_0)\}.$$

Then V is a neighbourhood of x_0, since $\{P(p_\alpha)\}_{\alpha \in A_n}$ is locally finite. Now let $x \in V$. Then

$$A_n(x_0) \subset A_n(x) \subset \bar{A}_n(x) \subset \bar{A}_n(x_0).$$

So if $\alpha_1, \ldots, \alpha_m$ are the elements of $\bar{A}_n(x)$ in order, we have, by the remark following (1) above,

$$f_n(x) = k_m((y_{\alpha_1}(x), \ldots, y_{\alpha_m}(x)), (p_{\alpha_1}(x), \ldots, p_{\alpha_m}(x))),$$
$$f_n(x_0) = k_m((y_{\alpha_1}(x_0), \ldots, y_{\alpha_m}(x_0)), (p_{\alpha_1}(x_0), \ldots, p_{\alpha_m}(x_0))).$$

Let

$$h_n(x, x_0) = k_m((y_{\alpha_1}(x_0), \ldots, y_{\alpha_m}(x_0)), (p_{\alpha_1}(x), \ldots, p_{\alpha_m}(x))).$$

Clearly

$$\rho(f_n(x), f_n(x_0)) \leqslant \rho(f_n(x), h_n(x, x_0)) + \rho(h_n(x, x_0), f_n(x_0)). \tag{2}$$

Now

$$x, x_0 \in P(p_{\alpha_i})^- \ (i = 1, \ldots, m),$$

whence

$$\sigma_n(y_{\alpha_i}(x), y_{\alpha_i}(x_0)) < 1$$

by 2.4(a), and hence, by 1.1(e′) at the beginning of this section,

$$\rho(f_n(x), h_n(x, x_0)) < 2^{-n}. \tag{3}$$

On the other hand, it follows from Definition 1.1(d) and the continuity of the functions $p_{\alpha_1}, \ldots, p_{\alpha_m}$ that to every $\epsilon > 0$ there corresponds a neighbourhood W of x_0 such that

$$\rho(h_n(x, x_0), f_n(x_0)) < \epsilon \qquad \text{if } x \in W. \tag{4}$$

Combining (2), (3), and (4), we conclude

$$\rho(f_n(x), f_n(x_0)) < 2^{-n} + \epsilon \qquad \text{if } x \in V \cap W,$$

and hence f_n is 2^{-n}-continuous at x_0.

(c) Pick a fixed n and $x \in X$. Letting $\alpha_1, \ldots, \alpha_m$ be the elements of $A_n(x)$ in order, we have

$$f_n(x) = k_m((y_{\alpha_1}(x), \ldots, y_{\alpha_m}(x)), (p_{\alpha_1}(x), \ldots, p_{\alpha_m}(x)));$$

letting $\gamma_1, \ldots, \gamma_m$ be the elements of $A_{n+1}(x)$ in order, we have

$$f_{n+1}(x) = k_{m'}((y_{\gamma_1}(x), \ldots, y_{\gamma_{m'}}(x)), (p_{\gamma_1}(x), \ldots, p_{\gamma_{m'}}(x))).$$

Now let

$$h_{n+1}(x) = k_{m'}((y_{\pi_n(\gamma_1)}(x), \ldots, y_{\pi_n(\gamma_{m'})}(x)), (p_{\gamma_1}(x), \ldots, p_{\gamma_{m'}}(x))).$$

To prove (c), it will suffice to show that

$$f_n(x) = h_{n+1}(x), \tag{5}$$

$$\rho(f_{n+1}(x), h_{n+1}(x)) < 2^{-n}. \tag{6}$$

Let us prove (5). From lemma 2.3 (c) we see that, for each $i \leqslant m'$, $\pi_n(\gamma_i) = \alpha_j$ for some $j \leqslant m$. We can, therefore, define a map $\pi: \{1, \ldots, m'\} \to \{1, \ldots, m\}$ by

$$\pi(i) = j \qquad \text{if, and only if,} \qquad \pi_n(\gamma_i) = \alpha_j.$$

It follows, again from Lemma 2.3(c), that

$$p_{\alpha_j}(x) = \sum_{i \in \pi} - 1_{(j)} p_{\gamma_i}(x) \qquad j = 1, \ldots, m.$$

Remembering that π_n, and hence π, is order-preserving, we now use Definition 1.1 (c′) (at the beginning of this section), with

$$y_j = y_{\alpha_j}(x), t_j = p_{\alpha_j}(x), y_i' = y_{\pi_n(\gamma_i)}(x),$$

and

$$t_i = p_{\gamma_i}(x) \ (j = 1, \ldots, m; i = 1, \ldots, m'),$$

to conclude (5).

It remains to establish (6). But this is easy since

$$\sigma_n(y_{\gamma_j}(x), y_{\pi(\gamma_j)}(x) < 1 \qquad j = 1, \ldots, m'$$

by Lemma 2.3 (b), and hence (6) follows from Definition 1.1 (e′) at the beginning of this section. This completes the proof of the lemma.

Theorem 1.3 is now a simple consequence of Lemma 3.1 and Lemma 2.2. In fact, the sequence

$$\{f_n\}_{n=1}^{\infty}$$

in Lemma 3.1 is uniformly Cauchy by 3.1 (c), and hence, by the completeness of Y, converges uniformly to an $f: X \to Y$. This f is continuous by Lemma 3.1 (b) and Lemma 2.2, and Lemma 3.1 (a) implies that $f(x) \in (\text{conv}(\phi(x)))^-$ for every $x \in X$. Thus f has all the required properties.

4. Proof of Theorem 1.5 from Theorem 1.3. Let $X, A \subset X$, Y, $\mathfrak{S} \subset 2^Y$, $\phi: X \to \mathfrak{S}$, and $g: A \to Y$ be as in Theorem 1.5. Since all our assumptions remain true if Y is replaced by its completion, we may as well assume that Y is complete.

(a) Assume that every $S \in \mathcal{S}$ is convex. Define $\psi: X \to 2^Y$ by

$$\psi(x) = \{g(x)\} \quad \text{if} \quad x \in A$$
$$\psi(x) = \phi(x) \quad \text{if} \quad x \notin A. \tag{1}$$

This ψ is lower semi-continuous by **(13**, Example 1.3***)**, and $\psi(x)$ is convex and closed for each $x \in X$ by the assumptions on ϕ and the remark following Definition 1.1. But Theorem 13 implies that such a ψ must have a selection f, and this f is the required extension of g.

(b) Assume that $\mathcal{S}$ is equi-locally convex. Let $Y' = \cup \mathcal{S}$. If $\mathcal{B}$ is as in Definition 1.4, use **(11**, Corollary 3.2**)** to obtain a metric τ on Y', agreeing with the topology, such that the family of all τ-spheres of radius 1 in Y' is a refinement of $\mathcal{B}$. We thus have

(2) *If* $S \in \mathcal{S}$, *and* $A \subset S$ *has* τ*-diameter* < 1, *then* A *is admissible, and* $\text{conv}(A) \subset S$, *whence* $(\text{conv}(A))^- \subset \hat{S} = S$.

Now let E be any Banach space, with metric d, containing (Y', τ) isometrically **(8)**, let $h: X \to E$ be a continuous extension of g **(1**, Theorem 1.4**)**, and let

$$W = \{x \in X \mid d(h(x), \phi(x)) < \tfrac{1}{2}\}.$$

Then obviously $W \subset A$, and W is open by the lower semi-continuity of ϕ. Pick an open $U \subset X$ such that $A \subset U \subset \bar{U} \subset W$. Let ψ be as in (1) above, and define $\theta: \bar{U} \to 2^Y$ by[7]

$$\theta(x) = \psi(x) \cap S_{\frac{1}{2}}(h(x)).$$

This θ is lower semi-continuous by **(13**, Proposition 2.5**)**. Moreover, by (2) $\theta(x)$ is admissible and $[\text{conv}(\theta(x))]^- \subset \phi(x)$ for all $x \in X$, and by the remark following Definition 1, $\text{conv}(\theta(x)) = \{g(x)\}$ for all $x \in A$. Our conclusion, therefore, follows from Theorem 1.3, with X replaced by $\bar{U}$, and ϕ by θ.

5. Geodesic structures. In this section, we define geodesic structures, and show how they generate convex structures. The closed unit interval will always be denoted by I.

Definition 5.1. A geodesic structure on a metric space E with metric ρ is a function $k: M \times I \to E$ (where $M \subset E \times E$) satisfying the following conditions:

(a) If $(x, x) \in M$, then $k(x, x, t) = x$ for all $t \in I$.

(b) If $(x_1, x_2) \in M$, then $k(x_1, x_2, 0) = x_1$ and $k(x_1, x_2, 1) = x_2$.

(c) If $(x_1, x_2) \in M$, $t \in I$, and $(k(x_1, x_2, t), x_2) \in M$, then $k(k(x_1, x_2, t), x_2; s) = k(x_1, x_2, t + s(1 - t))$ for all $s \in I$.

(d) For all $(x_1, x_2) \in M$, the map $t \to k(x_1, x_2, t)$, from I to E, is continuous.

(e) For each $\epsilon > 0$ there exist neigbourhoods $W_\epsilon \subset N_\epsilon$ of the diagonal in $E \times E$ which are small[8] of order ϵ, such that if $(x_1, x_2) \in M$ and $(y_1, y_2) \in M$,

[7] We use $S_r(y)$ to denote the open r-sphere about y.

[8] A neighbourhood U of the diagonal is *small of order* ϵ if $(x, y) \in U$ implies $\rho(x, y) < \epsilon$.

then $(x_1, y_1) \in N_*$ and $(x_2, y_2) \in W_*$ implies that $(k(x_1, x_2, t), k(y_1, y_2, t)) \in N_*$ for all $t \in I$.

Notice that Definition 5.1 is quite similar to Definition 1.1 with $n = 2$, but that 5.1 (e) is distinctly stronger than 1.1 (e) with $n = 2$. This extra strength is needed to carry out the inductive proof of Proposition 5.3 below.

Definition 5.2. Let E be a metric space with a geodesic structure. A subset S of E is *geodesic* if, whenever $x_1, x_2 \in S$, then $(x_1, x_2) \in M$ and $k(x_1, x_2, t) \in S$ for all $t \in I$. A family $\mathfrak{S}$ of subsets of E is *equi-locally geodesic* if there exists an open covering $\mathfrak{A}$ of E, such that, if $S \in \mathfrak{S}$, $A \in \mathfrak{A}$, and $x_1, x_2 \in (S \cap A)$, then $(x_1, x_2) \in M$ and $k(x_1, x_2, t) \in S$ for all $t \in I$.

Note that any family of geodesic sets is equi-locally geodesic.

PROPOSITION 5.3. *If E is a metric space with geodesic structure (M, k), then there exists a convex structure $\{M_n, k_n\}_{n=1}^{\infty}$ on E such that*

(a) *Every geodesic set $S \subset E$ is convex.*

(b) *Every equi-locally geodesic family of subsets of E is equi-locally convex.*

Proof. We define the sets $M_n \subset E^n$ and functions $k_n: M_n \times P_n \to E$ inductively as follows. First, let

$$M_1 = E, \qquad k_1(x, 1) = 1.$$

Suppose M_n and k_n have been defined, and let us define M_{n+1} and k_{n+1}. We introduce the following notation for the rest of this proof.

(1) If $x \in E^{n+1}$, then $\bar{x} \in E^n$ is defined by

$$\bar{x}_i = x_i \qquad (i = 1, \ldots, n).$$

(2) If $t \in P_{n+1}$ and $t_{n+1} \neq 1$, then $\bar{t} \in P_n$ is defined by

$$\bar{t}_i = \frac{t_i}{1 - t_{n+1}} \qquad (i = 1, \ldots, n).$$

Now let

(3) $M_{n+1} = \{x \in E^{n+1} | \bar{x} \in M_n, (k_n(\bar{x}, t), x_{n+1}) \in M \text{ for all } t \in P_n\}$,

(4) $k_{n+1}(x, t) = \begin{cases} x_{n+1} & \text{if } t_{n+1} = 1 \\ k(k_n(\bar{x}, \bar{t}), x_{n+1}, t_{n+1}) & \text{if } t_{n+1} \neq 1. \end{cases}$

Let us check that the conditions of Definition 1.1 are satisfied.

1.1 (a). This follows from our definition of k_1.

1.1 (b). For $n = 2$, this follows from the definitions. Suppose that it holds for n, and let us prove it for $n + 1$. Let $x \in M_{n+1}$, $i \leqslant n + 1$, and $t \in M_{n+1}$ with $t_i = 0$. If $i = n + 1$, then $\partial_i x = \bar{x}$, which is in M_n by (3), and $k_{n+1}(x, t) = k(k_n(\bar{x}, \bar{t}), x_{n+1}, 0) = k_n(\bar{x}, \bar{t}) = k_n(\partial_i x, \partial_i t)$, which is what we had to show.

Suppose, then, that $i \neq n + 1$, and let us first show that $\partial_i x \in M_n$. By (3), we must show that $(\partial_i x)^{\sim} \in M_{n-1}$, and that $(k_{n-1}((\partial_i x)^{\sim}, s), x_{n+1}) \in M$ for all $s \in P_{n-1}$. Now $(\partial_i x)^{\sim} = \partial_i \bar{x}$, and this is in M_{n-1} by the inductive hypothesis. As for the other requirement, pick an $s' \in P_n$ such that $s_i' = 0$ and

$\partial_i s' = s$; then $(k_{n-1}((\partial_i x)^\sim, s), x_{n+1}) = (k_{n-1}(\partial_i \bar{x}, \partial_i s'), x_{n+1}) = (k_n(\bar{x}, s'), x_{n+1})$ by the inductive hypothesis, and this last pair is in M by (3). Finally, let $t \in M_{n+1}$ with $t_i = 0$. If $t_{n+1} = 1$, then both sides of the equation in 1.1 (b) are x_{n+1}, and hence the equation is satisfied. If $t_{n+1} \neq 1$, then (by the induction hypothesis)

$$\begin{aligned} k_{n+1}(x, t) &= k(k_n(\bar{x}, \bar{t}), x_{n+1}, t_{n+1}) \\ &= k(k_{n-1}(\partial_i \bar{x}, \partial_i \bar{t}), x_{n+1}, t_{n+1}) \\ &= k(k_{n-1}((\partial_i x)^\sim, (\partial_i t)^\sim, x_{n+1}, t_{n+1}) \\ &= k_n(\partial_i x, \partial_i t), \end{aligned}$$

which is what we had to show.

1.1 (c). In the presence of 1.1 (b), 1.1 (c) is equivalent to the following condition, which we are going to verify: *If $x \in M_n$ $(n \geqslant 2)$ with $x_i = x_{i+1}$ for some $i < n$, and if $t, t' \in P_n$ with $t_j = t_j'$ for $j \neq i, i+1$, then $k_n(x, t) = k_n(x, t')$.*

For $n = 2$, this follows from 5.1 (a) and the definition of k_1. Suppose it is true for n, and let us prove it for $n + 1$.

Case 1. $i < n$. In this case, $t_{n+1} = t_{n+1}'$. If $t_{n+1} = 1$, then $k_{n+1}(x, t) = k_{n+1}(x, t') = x_{n+1}$. If $t_{n+1} \neq 1$, it suffices to show that $k_n(\bar{x}, \bar{t}) = k_n(\bar{x}, \bar{t}')$. But $\bar{t}_j = \bar{t}_j'$ for $j \neq i, i+1$, so this follows from the inductive hypothesis.

Case 2. $i = n$. In this case, $x_n = x_{n+1}$, and

$$\sum_{j=1}^{n-1} t_j = \sum_{j=1}^{n-1} t_j'.$$

If this sum is 0, then $k_{n+1}(x, t) = k_2((x_n, x_n), (t_n, t_{n+1}))$ (by 1.1 (b)) $= k(x_n, x_n, t_{n+1})$ (by (4)) $= x_n$ (by (5.1); similarly $k_{n+1}(x, t') = x_n$. If this sum is not zero, then $t_n + t_{n+1} = t_n' + t_{n+1}' < 1$, and $\bar{t} = \bar{t}'$. Applying (4) twice we get

$$\begin{aligned} k_{n+1}(x, t) &= k(k(k_{n-1}(\bar{\bar{x}}, \bar{\bar{t}}), x_n, t_n(1 - t_{n+1})^{-1}), x_n, t_{n+1}) \\ &= k(k_{n-1}(\bar{\bar{x}}, \bar{\bar{t}}), x_n, t_n + t_{n+1}) \quad \text{by 5.1 (c).} \end{aligned}$$

Similarly

$$k_{n+1}(x, t') = k(k_{n-1}(\bar{\bar{x}}, \bar{\bar{t}}'), x_n, t_n' + t_{n+1}').$$

But $t_n + t_{n+1} = t_n' + t_{n+1}'$, and $\bar{\bar{t}} = \bar{\bar{t}}'$, as already observed. Hence $k_{n+1}(x, t) = k_{n+1}(x, t')$, which is what had to be shown.

1.1 (e). Proof by induction. The result is clear for k_1. Suppose it is true for k_n, and let us prove it for k_{n+1}. Pick a fixed $y \in M_n$ and $s \in P_n$, and let us show that the map $t \to k_{n+1}(y, t)$, from P_{n+1} to E, is continuous at s.

Case 1. $s_{n+1} \neq 1$. In this case, for some neighbourhood U of s in P_{n+1}, $t \in U$ implies $t_{n+1} \neq 1$. Hence, for $t \in U$, $k_{n+1}(y, t)$ is given by the second formula in (4), and the required continuity follows from 5.1 (d) and our inductive hypothesis.

Case 2. $s_{n+1} = 1$. In this case $k_{n+1}(y, s) = y_{n+1}$. We will show that $k_{n+1}(y, t)$ is close to y_{n+1} if t_{n+1} is close to 1. To do this, let $A = \{k(y, t) | t \in P_n\}$ and, for each $\tau \in I$, define $f_\tau : A \to E$ by

$$f_\tau(a) = k(a, y_{n+1}, \tau).$$

Then

$$\begin{aligned} k_{n+1}(y, t) &= f_{t_{n+1}}(a_t) \quad \text{for some } a_t \in A \quad \text{if} \quad t_{n+1} \neq 1 \\ k_{n+1}(y, t) &= y_{n+1} \qquad\qquad\qquad\qquad\quad \text{if} \quad t_{n+1} = 1. \end{aligned}$$

It follows that it is sufficient to show that $f \to f_1$ uniformly as $\tau \to 1$. That $f_\tau \to f_1$ *pointwise* follows from 5.1 (d). Moreover A is compact, and $\{f_\tau\}_{\tau \in I}$ is equicontinuous by 5.1 (e). Hence (see, for instance, (**3**, p. 34, Prop. 14)), the topologies of pointwise convergence and uniform convergence coincide on $\{f_\tau\}_{\tau \in I}$, and hence $f_\tau \to f_1$ uniformly as $\tau \to 1$.

1.1 (f). For $\epsilon > 0$, let $V_\epsilon = W_\epsilon$, where W_ϵ is as in 5.1 (e). Let us show that this works. For k_1 this is clear. Suppose it works for k_n, and let us prove it for k_{n+1}. If $t_{n+1} = 1$, then $k_{n+1}(x, t) = x_{n+1}$ and $k_{n+1}(y, t) = y_{n+1}$, and we know that $\rho(x_{n+1}, y_{n+1}) < \epsilon$ since $(x_{n+1}, y_{n+1}) \in W_\epsilon$. If $t_{n+1} \neq 1$, then

$$\begin{aligned} k_{n+1}(x, t) &= k(k_n(\bar{x}, \bar{t}), x_{n+1}, t_{n+1}) \\ k_{n+1}(y, t) &= k(k_n(\bar{y}, \bar{t}), y_{n+1}, t_{n+1}). \end{aligned} \tag{5}$$

Now, since $(x_i, y_i) \in W_\epsilon$ for $i = 1, \ldots, n+1$, we have $(\bar{x}_i, \bar{y}_i) \in W_\epsilon$ for $i = 1, \ldots, n$, and hence our inductive hypothesis yields

$$k_n(\bar{x}, \bar{t}), k_n(\bar{y}, \bar{t})) \in N_\epsilon.$$

But $(x_{n+1}, y_{n+1}) \in W_\epsilon$ by assumption, and hence, by (5) and 5.1(e),

$$(k_{n+1}(x, t), k_{n+1}(y, t)) \in N_\epsilon,$$

and now the desired conclusion follows from the fact that N_ϵ is small of order ϵ.

We have now completed the proof that the sets M_n and functions k_n defined in (3) and (4) are a convex structure on E. To complete the proof, we must check conditions (a) and (b) of our proposition.

5.2 (a). Let $S \subset E$ be geodesic, and let $x \in S^n$. We must show that $x \in M_n$, and that $k_n(x, t) \in S$ for all $t \in P_n$. This is clear for $n = 1$. Let us therefore assume it for n, and prove it for $n + 1$. Now observe that, by the inductive hypothesis, $\bar{x} \in M_n$ and $k(\bar{x}, s) \in S$ for all $s \in P_n$. Since S is convex, it follows that $x \in M_{n+1}$ by (3), and $k_{n+1}(x, t) \in S$ for all $t \in P_{n0+1}$ by (4).

5.2 (b). Let $\mathfrak{S}$ be an equi-locally geodesic family of subsets of E, and let $\mathfrak{A}$ be as in definition of this concept in Definition 5.2. For each $x \in E$, pick an $A_x \in \mathfrak{A}$ which contains x, and then pick $r(x) > 0$ such that $S_{r(x)}(x) \subset A_x$. Let $B_x = W_{r(x)}(x)$, with W as in 5.1 (e), and let $\mathfrak{B} = \{B_x\}_{x \in E}$. Let us show that this $\mathfrak{B}$ works.

Let $S \in \mathfrak{S}$, $x_0 \in E$, and let $T = S \cap B_{x_0}$. We must show that if $x \in T^n$,

then $x \in M_n$, and $k_n(x, t) \in S$ for all $t \in P_n$. We will prove this inductively and, to keep the induction going, we will also show that, for all n,

$$(k_n(x, t), x_0) \in N_{r(x_0)} \quad \text{for all} \quad t \in P_n, \tag{6}$$

where $N_{r(x_0)}$ is as in 5.1(e). Now for $n = 1$, all this is clear. Suppose it is true for n, and let us prove it for $n + 1$.

By the inductive hypothesis, $\bar{x} \in M_n$ and

$$k_n(\bar{x}, s) \in (S \cap A_{x_0})$$

for all $s \in P_n$. Since also

$$x_{n+1} \in (S \cap A_{x_0}),$$

it follows that $x \in M_{n+1}$ by (3), and $k_{n+1}(x, t) \in S$ for all $t \in P_{n+1}$ by (4). It remains to check (6) for $n + 1$. If $t_{n+1} = 1$, this is clear, since then $k_n(x, t) = x_{n+1}$. If $t_{n+1} \neq 1$, note that

$$(k_n(\bar{x}, \bar{t}), x_0) \in N_{r(x_0)}$$

by the inductive hypothesis, and

$$(x_{n+1}, x_0) \in W_{r(x_0)}$$

by assumption. Hence

$$(k_n(x, t), x_0) = (k(k_n(\bar{x}, \bar{t}), x_{n+1}, t_{n+1}), x_0) \in N_{r(x_0)}$$

by (4) and 5.1 (c), and that is what had to be shown. This completes the proof.

To conclude this section, let us record the following consequence of Proposition 5.3.

THEOREM 5.4. *Theorem* 1.5 *remains true if "(locally) convex" is replaced by "(locally) geodesic".*

6. Riemannian manifolds. Let E be a Riemannian manifold with Riemannian metric ρ. If $x_1, x_2 \in X$ have a unique shortest geodesic joining them, this geodesic is called a *segment*. Let

$$L = \{(x_1, x_2) \in E \times E | x_1 \text{ and } x_2 \text{ are joined by a segment}\}. \tag{1}$$

If $(x_1, x_2) \in L$, then the segment from x_1 to x_2 is given by a continuous

$$g_{x_1, x_2}: I \to E.$$

Define $h: L \times I \to E$ by

$$h(x_1, x_2, t) = g_{x_1, x_2}(t). \tag{2}$$

The elementary properties of geodesics imply that (L, h) satisfies conditions (a) – (d) of Definition 5.1. In general, of course, 5.1 (e) is not satisfied. However we have the following theorem of Nijenhuis **(17)**.

THEOREM. *Every $p \in E$ has a spherical neighbourhood $V(p)$ with the following property:*

(a) *If $x_1, x_2 \in V(p)$, then $(x_1, x_2) \in L$, and $h(x_1, x_2, t) \in V(p)$ for all $t \in I$.*

(b) *If $\epsilon > 0$, and $x_1, x_2, y_1, y_2 \in V(p)$ with $\rho(x_1, y_1) < \epsilon$ and $\rho(x_2, y_2) < \frac{1}{2}\epsilon$, then $\rho(h(x_1, x_2, t), h(y_1, y_2, t)) < \epsilon$ for all $t \in I$.*

It follows immediately from the above theorem that, if $M = V(p) \times V(p)$ and $k = h|M$, then (M, h), in addition to satisfying conditions (a) – (d) of Definition 5.1, also satisfies condition (e) (by taking N_ϵ to be the ϵ-neighbourhood of the diagonal, and W_ϵ the $\frac{1}{2}\epsilon$-neighbourhood), and hence we conclude

PROPOSITION 6.1. *Every point p of a Riemannian manifold E has a spherical neighbourhood $V(p)$ such that every pair of points in $V(p)$ is joined by a unique geodesic segment, and these geodesic segments generate a geodesic structure on $V(p)$ under which $V(p)$ is a geodesic set.*

While the above proposition is sufficient for our application to Lie groups in the next section, let us conclude this section by proving the following more "global" result.

PROPOSITION 6.2. *On every compact Riemannian manifold E there exists a geodesic structure (M, k) such that*

(a) *If L and h are as in (1) and (2), then $M \subset L$ and $k = h|M$.*

(b) *Every $p \in E$ has a neighbourhood $U(p)$ such that $U(p) \times U(p) \subset M$.*

Proof. For each $p \in E$, let $S(p)$ be the open sphere about p whose radius is half the radius of the sphere $V(p)$ in Nijenhuis's theorem. Then $\{S(p)\}_{p \in E}$ is an open covering of E, and by compactness there exsits a finite subcovering $\{S(p_i)\}_{i=1}^{n}$. With L and h as in (1) and (2), let

$$\begin{aligned} M &= \{(x, y) \in L | x, y \in S(p_i) \quad \text{for some } i = 1, \ldots, n\} \\ k &= h|M \end{aligned} \tag{3}$$

The only requirement that needs further checking is that (M, k) satisfies condition (e) of Definition 5.1.

Let $\epsilon > 0$, and let us define the required neighbourhoods W_ϵ and N_ϵ of the diagonal. For each i, let r_i be the radius of $S(p_i)$, and let

$$\gamma = \min(\min\{r_i | i = 1, \ldots, n\}, \epsilon).$$

Now let

$$\begin{aligned} N_\epsilon &= \{(x, y) \in E \times E | \rho(x, y) < \gamma\}, \\ W_\epsilon &= \{(x, y) \in E \times E | \rho(x, y) < \tfrac{1}{2}\gamma\}. \end{aligned} \tag{4}$$

To see that this works, suppose that $(x_1, x_2) \in M$, $(y_1, y_2) \in M$, and that $(x_1, y_1) \in N_\epsilon$ and $(x_2, y_2) \in W_\epsilon$. Since $(x_1, x_2) \in M$, it follows from (3) that there exists a positive integer $k \leq n$ such that

$$x_1, x_2 \in S(p_k) \subset V(p_k).$$

Since $\rho(x_i, y_i) < \gamma < r_k$ for $i = 1, 2$, it follows that

$$y_1, y_2 \in V(p_k).$$

Finally, observe that $\rho(x_1, y_1) < \gamma$, while $(x_2, y_2) < \frac{1}{2}\gamma$. We therefore apply Nijenhuis's theorem to $V(p_k)$ to conclude that, for all $t \in I$, $\rho(k(x_1, y_1, t), k(x_2, y_2, t)) < \gamma < \epsilon$, which is what we had to show.

7. Locally convex groups. The reader is reminded that, as elsewhere in this paper, A^n always (even in a group) denotes the n-fold Cartesian product of A for any positive integer n. Moreover, if G is a group, $x \in G^n$, and $z \in G$, then xz will denote $(x_1z, \ldots, x_nz) \in G^n$, while if $K \subset G^n$ and $z \in G$, then Kz will denote $\{xz | x \in K\} \subset G^n$.

Definition 7.1. A metrizable topological group H with right-invariant metric d is *locally convex* if there exists a convex structure $\{L_n, h_n\}_{n=1}^{\infty}$ on H, and a neighbourhood W of the identity, such that the following conditions are satisfied for all n.

(a) $L_n = W^n$.

(b) h_n is right-invariant; that is, if $x \in W^n$, $z \in H$, and $xz \in W^n$, then $h_n(xz, t) = (h_n(x, t))z$ for all $t \in P_n$.

(c) The left and right uniform structures on H coincide on W.

(d) Condition (e) of Definition 1.1 is satisfied with V_ϵ of the special form $\{(x, y) \in E \times E | d(x, y) < \delta(\epsilon)\}$ for some $\delta(\epsilon) > 0$.

If we can take $W = H$, then H is called a *convex* group.

Note that this definition does not depend on which right-invariant metric is used, since any two such metrics are uniformly equivalent.

The following are some examples of locally convex groups.

(1) A locally convex metrizable topological linear space, with the usual convex structure. (This group is, in fact, convex.)

(2) A Lie group. A Lie group can be made into a Riemannian manifold with right-invariant Riemannian metric. By Proposition 6.1, there exists a neighbourhood U of e on which the segments generate a geodesic structure (M, k) making U into a geodesic set. Since the metric is right-invariant, so is k. By Proposition 5.3, this geodesic structure generates a convex structure $\{M_n, k_n\}_{n=1}^{\infty}$ on U which makes U a convex set; hence $M_n = U^n$. From the way k_n is defined in terms of k in the proof of Proposition 5.3, it follows that each k_n is right-invariant along with k. Now let W be a compact subneighbourhood of U, and let $h_n = k_n | W^n$. Then all requirements are satisfied, the compactness of W taking care of 7.1 (c) and (d).

(3) The multiplicative group H_0 of invertible elements in a Banach algebra B with unit e. Let W be the sphere of radius $\frac{1}{2}$ about e. For the convex structure, use the ordinary linear one. The right and left uniform structures on H coincide on W with the uniform structure induced by the norm. More generally, one can take any group of the form $H_0 \cap (I + e)$, where I is an ideal in B.

(4) A topological group which is locally isomorphic to a locally convex group.

(5) A finite Cartesian product of locally convex groups.

THEOREM 7.2. *If G is a metrizable group, and H a closed, complete (locally), convex subgroup, then there exists a (local) cross-section.*[9]

Proof. Let ρ be a *left* invariant metric on G. We shall construct a convex structure $\{M_n, k_n\}_{n=1}^{\infty}$ on (G, ρ) such that the collection $\mathfrak{H}$ of *right* cosets of H becomes equi-locally convex (and such that each right coset of H becomes a convex set in case H is convex.) The theorem then follows from Theorem 1.5, which applies because the natural map $u: G \to G/H$ is open, and hence the carrier $\phi: G/H \to 2^G$, defined by $\phi(x) = u^{-1}(x)$, is lower semi-continuous **(13,** Example 1.3[*]**)**.

By 1.1 (e) and the remark following Definition 1.1, there exists a neighbourhood W_1 of the identity e such that, whenever $x \in W_1^n$, then $h_n(x, t) \in W$ for all $t \in P_n$. Pick a symmetric neighbourhood W_2 of e such that $W_2 \cdot W_2 \subset W_1$ Now let

$$M_n' = \cup\{W_1^n z | z \in G\},$$
$$M_n = \cup\{W_2^n z | z \in G\}.$$

Define $k_n': M_n' \times P_n \to G$ by

$$k_n'(x, t) = (h_n(xz^{-1}, t))\, z,$$

where z is any element of G such that $x \in W_1^n z$, and let

$$k_n = k_n' | (M_n \times P_n).$$

It follows from Definition 7.1 (b) that the definition of k_n', and hence that of k_n, does not depend on the choice of z. That $\{M_n, k_n\}_{n=1}^{\infty}$ satisfies conditions (a) – (d) of Definition 1.1 follows from the fact that $\{L_n, h_n\}_{n=1}^{\infty}$ satisfies them. It remains to verify that $\{M_n, k_n\}_{n=1}^{\infty}$ satisfies condition (e) of Definition 1.1.

Observe first that, for any topological group G and any $z \in G$, right translation by z ($z \to xz$) is a uniformity isomorphism for both the right and left uniform structures on G. For the right one this is clear. For the left one, it follows from the fact that our map is a composition of $x \to z^{-1}xz$ and $y \to zy$; the first of these is a group isomorphism, and the second one is a left translation.

It follows from the previous paragraph, applied to the left uniform structure, that 7.1(c) and (d) imply

(1) *To every $z \in G$ and $\epsilon > 0$ corresponds a $\delta(\epsilon, z) > 0$ such that, for all n, if $x, y \in W_1^n z$ and $\rho(x_i, y_i) < \delta(\epsilon, z)$ for $i = 1, \ldots, n$, then $\rho(k_n'(x, t), k_n'(y, t) < \epsilon$ for all $t \in P_n$.*

[9] See footnote 2, p. 553.

Now for each $\epsilon > 0$ and $z \in G$, let

(2) $\xi(\epsilon, z) = \frac{1}{4}\delta(\frac{1}{2}\epsilon, z)$.

(3) $V_\epsilon = \{(x, y) \in G \times G | x, y \in S_{\xi(\epsilon,z)}(z) \text{ for some } z \in G\}$.

Let us check that this V_ϵ works. We must show that, if $x, y \in M_n$, and if for each $i = 1, \ldots, n$ there exists a $z_i \in G$ such that

$$x_i, y_i \in S_{\xi(\epsilon,z_i)}(z_i),$$

then

(4) $\rho(k_n(x, t), k_n(y, t)) < \epsilon$

for all $t \in P_n$.

Pick $k \leqslant n$ such that $\xi(\epsilon, z_k) = \max_{i=1}^{n} \xi(\epsilon, z_i)$, and let $a = x_k^{-1}z_k$, $b = y_k^{-1}z_k$. Note that, since ρ is left invariant, $\rho(g, ga) = \rho(x_k, z_k) < \xi(\epsilon, z_k)$ and $\rho(g, gb) = \rho(y_k, z_k) < \xi(\epsilon, z_k)$ for all $g \in G$. For $i = 1, \ldots, n$, let $x_i' = x_i a$, $y_i' = y_i b$. Then

$$\rho(x_i', y_i') \leqslant \rho(x_i', x_i) + \rho(x_i, y_i) + \rho(y_i, y_i') \leqslant 4\xi(\epsilon, z_k) = \delta(\tfrac{1}{2}\epsilon, z_k).$$

Now $x \in M_n$, so $x \in W_2^n r$ for some $r \in G$, and hence $x' \in W_2^n ra$. Thus, for $i = 1, \ldots, n$, $x_i' \in W_2 ra = W_2(rx_k^{-1})z_k \subset (W_2 \cdot W_2)z_k \subset W_1 z_k$; in other words, $x' \in W_1^n z_k$. Similarly, $y' \in W_1^n z_k$. We can therefore, apply (1), with $z = z_k$, to conclude that $\rho(k_n'(x', t), k_n'(y', t)) < \frac{1}{2}\epsilon$ for all $t \in P_n$.

Now note that, for all $t \in P_n$,

$$k_n'(x', t) = h_n(x'z_k^{-1}, t)\, z_k = h_n(xx_k^{-1}, t)\, x_k a = k_n(x, t)a,$$

and hence $\rho(k'_n(x', t), k_n(x, t)) < \xi(\epsilon, z_k)$. Similarly $\rho(k_n'(y', t), k_n(y, t)) < \xi(\epsilon, z_k)$ for all $t \in P_n$. Hence, for all $t \in P_n$,

$$\rho(k_n(x, t), k_n(y, t)) \leqslant \rho(k_n(x, t), k_n'(x', t)) + \rho(k_n'(x', t), k_n'(y', t)) + \rho(k_n'(y', t), k_n(y, t))$$

which is what we had to show.

To complete the proof, we must show that the family $\mathfrak{H}$ of right cosets of H is equi-locally convex, and that each right coset of H is a convex set in case H is convex. The latter assertion is clear. To prove the former, let d be a *right* invariant metric on G, and pick $r > 0$ such that the d-sphere of radius r about the identity e in H is contained in W_2. Since $x \in M_n$ if, and only if, $x_1, \ldots, x_2$ are all in some right translate of W_2, it follows that $x \in M_n$ whenever $x_1, \ldots, x_n$ is a subset of some member of $\mathfrak{H}$ having d-diameter $< r$. Hence for $\mathfrak{B}$ we simply pick the family of d-spheres of radius $\frac{1}{2}r$ with centres in G.

Combining Theorem 7.2 with Example (2) yields a new proof—at least in metrizable case—of a cross-section theorem of A. Gleason. When combined with Examples (1) and (4), Theorem 7.2 yields

COROLLARY 7.3. *Let G be a metrizable group, and H a closed subgroup which is (locally) isomorphic to the additive group of a complete, metrizable, locally convex topological linear space. Then there exists a (local) cross section.*[10]

This corollary generalizes a result of Bartle and Graves **(2**, Theorem 4**)** (see also **(13**, p. 364**)**), where G is assumed to be a Banach space and H a closed linear subspace.

8. Homotopy extension, covering homotopy and fibre spaces. A common feature of both the homotopy extension theorem **(6**, Theorem VI, 5**)** and the covering homotopy theorem **(19**, Theorem 11.7**)** is that, speaking very roughly, special conditions on the domain permit one to obtain a continuous function globally even though the range is well behaved only locally. In this section we obtain a result (Theorem 8.3) on continuous selections which seems to incorporate the essential aspects of both these theorems. For finite dimensional domains, this was already done in **(15**, Theorem 3.4**)**, and both statement and proof of Theorem 8.3 parallel those of its finite-dimensional analogue. After developing the necessary preliminary concepts, we will thus be able to dispose of the proof of Theorem 8.3 simply by pointing out the obvious modifications which must be made in the proof of **(15**, Theorem 3.4**)**.

Definition 8.1. Let E be a metric space with convex structure. Then a family $\mathfrak{S}$ of subsets of E is *uniformly equi-locally convex* if there exists an $r > 0$ such that, whenever $S \in \mathfrak{S}$ and $A \subset S$ has diameter $< r$, then A is admissible and $\operatorname{conv}(A) \subset S$.

Note that any uniformly equi-locally convex family of sets is equi-locally convex.

Using Theorem 1.3, we now prove

PROPOSITION 8.2. *Let (Y, ρ) be a metric space with convex structure, $\mathfrak{S}$ a uniformly equi-locally convex family of complete, non-empty subsets of Y, and let $r > 0$ be as in Definition* 8.1. *Let X be paracompact, $\phi: X \to \mathfrak{S}$ lower semi-continuous, and suppose there exists a continuous $g: X \to Y$ such that $\rho(g(x), \phi(x)) < r$ for every $x \in X$. Then there exists a selection for ϕ.*

Proof. Since all assumptions remain unchanged if Y is replaced by its completion, we may assume that Y is complete. Define $\psi: X \to 2^Y$ by

$$\psi(x) = \phi(x) \cap S_r(g(x)).$$

Then ψ is lower semi-continuous by **(13**, Proposition 2.5**)** and, for every $x \in X$, $\psi(x)$ is admissible and $\operatorname{conv}(\psi(x)) \subset \phi(x)$. Hence Theorem 1.3 asserts the existence of a continuous $f: X \to Y$ such that, for every $x \in X$,

$$f(x) \in [\operatorname{conv}(\psi(x))]^- \subset \phi(x).$$

This f is the required selection, and the proof is complete.

[10]See footnote 2, p. 556.

The previous theorems in this paper only required the carrier $\phi: X \to 2^Y$ to be lower semi-continuous. In the following theorem, however, just as in [15], we require it to be *continuous*; that is, given $\epsilon > 0$, every $x_0 \in X$ has a neighbourhood U such that

$$\phi(x_0) \subset S_\epsilon(\phi(x)), \quad \phi(x) \subset S_\epsilon(\phi(x_0))$$

for every $x \in U$.

The principal result of this section can now be stated as follows.

THEOREM 8.3. *Let Y be a metric space with convex structure, and $\mathfrak{S}$ a uniformly equi-locally convex family of complete non-empty subsets of Y. Let Z be paracompact, $X = Z \times I$, and $\phi: X \to \mathfrak{S}$ continuous. Finally, let $B \subset Z$ be closed, and define $A \subset X$ by*

$$A = (Z \times \{0\}) \cup (B \times I).$$

Then every selection for $\phi|A$ can be extended to a selection for ϕ.

Proof. This theorem is identical with **(15**, Theorem 3.4**)** except that no dimensional restrictions are placed on Z, and that $\mathfrak{S}$ is uniformly equi-locally convex instead of uniformly equi-LC^n. The proof of our theorem is similarly identical with the first part of the proof of **(15**, Theorem 3.4**)** (the second part takes care of a dimensional difficulty), provided "uniformly equi-LC^n" is replaced by "uniformly equi-locally convex," and the reference to **(14**, Theorem 9.1**)** in the proof of Lemma 3.3 of **(15)** is replaced by a reference to our Proposition 8.2. This is all that need be said, and the proof is thus complete.

Before continuing, let us observe that if every $S \in \mathfrak{S}$ were actually *convex* in Theorem 8.3, then this theorem would be a special case of Theorem 1.5(a). Note also that if ϕ is a constant map, then Theorem 8.3 is simply a homotopy extension theorem, while if B is empty, we get a theorem of the covering-homotopy type.

As mentioned at the beginning of this section, Theorem 8.3 is valid because of the special relation of X to $A \subset X$. However, the relation need not be quite as special as all that. In fact, it is an easy consequence of Theorem 8.3 that it is sufficient to assume that X is paracompact, and that A is a *generalized deformation retract* of X in the following sense: There exists a continuous $r: X \times I \to A$ such that $r(x, 0) = x$ if $x \in X$, $r(x, 1) \in A$ if $x \in X$, and $r(x, t) \in A$ if $x \in A$ and $t \in I$. The simple proof of this generalization from Theorem 8.3 can be omitted, since it is identical with the proof of how Theorem 6.2 follows from Lemma 5.3 in **(15)**.

Just as **(15**, Theorem 3.4**)** was applicable to fibre spaces in the sense of Serre **(18)**, Theorem 8.3 is applicable to fibre spaces in the sense of Hurewicz **(7)**. These are, by definition, triples (E, p, B), where E and B are topological spaces with $p: E \to B$ continuous and onto, satisfying the following condition: If Z is paracompact, $k: Z \times I \to B$ continuous, and if $\phi: Z \times I \to 2^E$ is defined by $\phi(z, t) = p^{-1}(k(z, t))$, then every selection for $\phi|Z \times \{0\}$ can

be extended to a selection for ϕ. The following result is now an immediate consequence of Theorem 8.3.

COROLLARY 8.4. *Let E, B be metric spaces, $p: E \to B$ continuous and onto, and suppose that*

(a) *each $p^{-1}(x)$ is complete,*

(b) *$\{p^{-1}(x)\}_{x \in B}$ is uniformly equi-locally convex,*

(c) *the carrier $\psi: B \to 2^E$, defined by $\psi(x) = p^{-1}(x)$, is continuous.*

Then (E, p, B) is a fibre space in the sense of Hurewicz (7).

REFERENCES

1. R. Arens, *Extensions of functions in fully normal spaces*, Pacific J. Math., *2* (1952), 11–23.
2. R. Bartle and L. M. Graves, *Mappings between function spaces*, Trans. Amer. Math. Soc., *72* (1952), 400–13.
3. N. Bourbaki, *Topologie générale* (Paris, 1949), chapter X.
4. J. Dieudonné, *Une généralization des espaces compacts*, J. de Math. pur. et appl., *23* (1944), 65–76.
5. J. Dugundji, *An extension of Tietze's theorem*, Pacific J. Math., *1* (1951), 175–86.
6. W. Hurewicz and H. Wallman, *Dimension theory* (Princeton, 1948).
7. W. Hurewicz, *On the content of fibre space*, Proc. Nat. Acad. Sci., *41* (1955), 956–61.
8. V. L. Klee, *Some characterizations of compactness*, Amer. Math. Monthly, *58* (1951), 389–93.
9. S. Mazur, *Ueber die kleinste convexe Menge, die eine gegebene kompakte Menge enthält*, Studia Math., *2* (1930), 7–9.
10. E. Michael, *Some extension theorems for continuous functions*, Pacific J. Math., *3* (1953), 798–806.
11. ———, *Local properties of topological spaces*, Duke Math. J., *21* (1954), 163–71.
12. ———, *Selected selection theorems*, Amer. Math. Monthly, *63* (1956), 233–8.
13. ———, *Continuous selections I*, Ann. Math., *63* (1956), 361–82.
14. ———, *Continuous selections II*, Ann. Math., *64* (1956), 562–80.
15. ———, *Continuous selections III*, Ann. Math., *65* (1957), 375–90.
16. ———, *A note on semi-continuous set-valued functions*, to appear.
17. A. Nijenhuis, *Local hyperconvexity of Riemannian manifolds*, to appear.
18. J. P. Serre, *Homologie singulière des éspaces fibrés*, Ann. Math., *54* (1951), 524–605.
19. N. Steenrod, *The topology of fibre bundles* (Princeton, 1951).

University of Washington

A NOTE ON HYPERCONVEXITY IN RIEMANNIAN MANIFOLDS

ALBERT NIJENHUIS

1. Summary. Let M denote a connected Riemannian manifold of class C^3, with positive definite C^2 metric. The curvature tensor then exists, and is continuous.

By a classical theorem of J. H. C. Whitehead **(1)**, every point x of M has the property that all sufficiently small spherical neighbourhoods V of x are *convex*; that is, (i) to every $y, z \in V$ there is one and only one geodesic segment yz in M which is the shortest path joining them: $f:[0, 1] \to M$, $f(0) = y$, $f(1) = z$; and (ii) this segment yz lies entirely in V: $f([0, 1]) \subset V$; (iii) if f is parametrized proportional to arc length, then $f(t)$ is a C^2 function of y, t, and z.

Let V be a convex set in M; and let $y_1, y_2, z_1, z_2 \in V$. Let f_1, f_2: $[0, 1] \to V$ denote the geodesic segments y_1z_1, y_2z_2, each parametrized proportional to arc length. Then for each t the points $f_1(t), f_2(t)$ are called *corresponding points* of the geodesic segments y_1z_1, y_2z_2. In particular, y_1, y_2 are corresponding points; and so are z_1, z_2.

The distance between points x and y, denoted by $\rho(x, y)$, is the greatest lower bound of the lengths of rectifiable paths joining x and y. The diameter $D(A)$ of a set A is, as usual

$$D(A) = \sup_{x,y \in A} \rho(x, y).$$

Let V be a convex open set of M, and γ a positive number. Then V is called *γ-hyperconvex* if for every positive number ϵ and any geodesic segments y_1z_1, y_2z_2 in V, the inequalities $\rho(y_1, y_2) < \epsilon$, $\rho(z_1, z_2) < \gamma\epsilon$ imply that corresponding points of y_1z_1, y_2z_2 have distance less than ϵ. Clearly, γ has to satisfy $\gamma \leqslant 1$.

If V is γ-hyperconvex, and W is a convex subset of V, then also W is γ-hyperconvex. If V is γ-hyperconvex, and if $0 < \gamma' < \gamma$, then V is also γ'-hyperconvex.

THEOREM I. *Every point of a Riemannian manifold has a $\frac{1}{2}$-hyperconvex neighbourhood.*

Received June 2, 1958. Research supported by an Office of Naval Research Contract at the University of Washington.

This theorem* is a corollary of Theorem II, which makes use of the concept of maximal curvature, defined as follows.

If u and v are tangent vectors at some $x \in M$, with components u^λ, v^λ ($\kappa, \lambda, \mu, \nu, \ldots, = 1, \ldots, n$) with respect to some local co-ordinates; and if $K_{\alpha\lambda\mu\nu}$ are the covariant components of the curvature tensor $K^{\cdots\nu}_{\alpha\lambda\mu}$, then $K(u, v)$ is defined as $-K_{\alpha\lambda\mu\nu}u^\alpha v^\lambda u^\mu v^\nu$. When u and v are perpendicular unit vectors, $K(u, v)$ is the *sectional curvature*, which clearly depends only on the 2-plane at x, spanned by u and v. If A is a subset of M, then $k(A)$, the *maximum curvature* of A, denotes the least upper bound of all numbers $K(u, v)$, with u, v being perpendicular unit vectors spanning all 2-planes at all points of A. $k(A)$ may be $+\infty$.

Our main result is the following theorem.

THEOREM II. *If the maximum curvature $k = k(V)$ of a convex open set V in M is non-positive: $k \leqslant 0$, then V is 1-hyperconvex. If $k > 0$, and if the diameter $D = D(V)$ is such that $kD^2 < \pi^2/4$, then V is γ-hyperconvex, where $\gamma = \cos Dk^{1/2}$.*

The case $k \leqslant 0$ is a much weakened formulation of results obtained by H. Busemann (**3**, Theorems (36.4) and (36.17)); the first of which states that the distance between corresponding points of geodesics in a G-space is a convex function of the linear parameter on the geodesics. Since, however, the case $k \leqslant 0$ is naturally included in our line of argument, the reader will find a new proof for this result.

Every sufficiently small spherical neighbourhood V of a point $x \in M$ is convex and has compact closure, which implies that $k(V)$ is finite. Since $k(V)$ is non-increasing when the radius of V tends to zero, it follows that $k(V)D(V)^2$ tends to zero as the radius of V approaches zero. This proves Theorem I, assuming Theorem II.

The proof of Theorem II is based on estimates of the solution of systems of linear differential equations whose prototype is $y'' = f(x)y$.

2. Estimates on certain systems of linear differential equations.

LEMMA 1. *Let $K(s)$ be a continuous family of linear transformations in E^n, with origin O, defined in the finite closed interval $0 \leqslant s \leqslant l$. Let $\mathbf{y}(s)$ be a solution of $\mathbf{y} + K(s)\mathbf{y} = 0$ in the interval $[0, l]$, and let S denote the unit sphere around O. Then there is a C^2 path $\mathbf{x}(s)$ on S such that for each $s \in [0, l]$, the points O, $\mathbf{x}(s)$, $\mathbf{y}(s)$ are collinear ($\mathbf{x}(s)$ is called a spherical image of $\mathbf{y}(s)$).*

*This statement was conjectured by E. A. Michael in a slightly weaker form obtained by replacing $\frac{1}{2}$-hyperconvexity of a convex set V by the following property: to every $\epsilon > 0$ there is a $\delta > 0$ such that if $y_1, y_2, z_1, z_2 \in V$, $\rho(y_1, y_2) < \epsilon$, $\rho(z_1, z_2) < \delta$, then the distance between corresponding points of y_1, z_1, y_2z_2 is less than ϵ. This conjecture was submitted by Dr. Michael to a number of mathematicians, including H. E. Rauch, who vouched for it, and L. W. Green, who obtained a written proof (spring, 1957) of the weaker result which one gets by permitting δ to be equal to zero. Theorem I, in its present form, has been applied in (**2**) to the theory of continuous selections.

Proof. If $\mathbf{y} \equiv 0$ the lemma is trivial, because any C^2 path $\mathbf{x}(s)$ on S is a spherical image. If $\mathbf{y}(s) \neq 0$ for all $s \in [0, l]$ the proof is simple because $\mathbf{x}(s) = \mathbf{y}(s) \cdot |\mathbf{y}(s)|^{-1}$ is a spherical image. Now assume that $\mathbf{y}(s)$ has the zeros $s_1, s_2, \ldots$, but $\mathbf{y} \not\equiv 0$. The number N of these s_i must be finite, because at any accumulation point s' one has

$$\mathbf{y}(s') = \dot{\mathbf{y}}(s') = 0;$$

hence $\mathbf{y} \equiv 0$. Consider

$$\mathbf{z}(s) = \mathbf{y}(s)(s - s_1)^{-1}(s - s_2)^{-1} \ldots (s - s_N)^{-1}$$

for $s \neq s_1, \ldots, s_N$, and

$$\mathbf{z}(s_i) = \lim_{s \to s_i} \mathbf{z}(s) = \dot{\mathbf{y}}(s_i) \prod_{j \neq i} (s_i - s_j)^{-1}.$$

Then $\mathbf{z}(s) \neq 0$ for all $s \in [0, l]$ because

$$\dot{\mathbf{y}}(s_i) \neq 0.$$

Every spherical image of $\mathbf{z}(s)$ is a spherical image of $\mathbf{y}(s)$; and thus the problem has been reduced to a previous case provided $\mathbf{z}(s)$ is of class C^2 for $s = s_1, \ldots, s_N$. In verifying this, the following simple application of the mean value theorem is helpful: "*If a function f is continuous in $[a, b]$, differentiable at all points of $[a, b]$ except some $c \in [a, b]$, and if*

$$\lim_{s \to c} f'(s) = L$$

exists, then $f'(c)$ exists and equals L; whence f' is continuous at c". To show that $\mathbf{z}(s)$ is of class C^2 at s_i it suffices to show this for $\mathbf{z}_i(s) = \mathbf{y}(s)(s - s_i)^{-1}$. We use de l'Hopital's Rule:

$$\lim_{s \to s_i} \dot{\mathbf{z}}_i(s) = \lim \frac{\dot{\mathbf{y}}(s)(s - s_i) - \mathbf{y}(s)}{(s - s_i)^2} = \lim \frac{\ddot{\mathbf{y}}(s)(s - s_i)}{2(s - s_i)} = -\tfrac{1}{2}K(s_i)\mathbf{y}(s_i) = 0.$$

Hence, $\dot{\mathbf{z}}_i(s_i) = 0$, and $\dot{\mathbf{z}}_i$ is continuous in a neighbourhood of s_i in $[0, l]$. The procedure is repeated for $\ddot{\mathbf{z}}_i$, and one thus finds $\ddot{\mathbf{z}}_i(s_i) = -\tfrac{1}{3}K(s_i)\dot{\mathbf{y}}(s_i)$; and $\ddot{\mathbf{z}}_i(s)$ is continuous at s_i. Hence, $\ddot{\mathbf{z}}(s)$ exists and is continuous in $[0, l]$.

LEMMA 2. *If $K(s)$ is a continuous family of linear transformations in E^n; $0 \leqslant s \leqslant l$, and m is an upper bound for the inner product $(\mathbf{u}, K(s)\mathbf{u})$ for all $s \in [0, l]$ and all unit vectors $\mathbf{u}$; if $\mathbf{y}(s)$ is a solution of*

$$\ddot{\mathbf{y}} + K(s)\mathbf{y} = 0,$$

and if $\mathbf{x}(s)$ is a spherical image of $\mathbf{y}(s)$; with $\mathbf{y}(s) = \lambda(s)\mathbf{x}(s)$; then λ satisfies a differential equation $\ddot{\lambda}(s) + \phi(s)\lambda(s) = 0$, where $\phi(s)$ is continuous in s; $0 \leqslant s \leqslant l$; and $\phi(s) \leqslant m$.

Proof. Since $\mathbf{y}(s)$ and $\mathbf{x}(s)$ are of class C^2, and $\mathbf{x}(s) \neq 0$; $\lambda(s)$ is also of class C^2. Substituting $\mathbf{y} = \lambda\mathbf{x}$ into $\mathbf{y} + K(s)\mathbf{y} = 0$ one finds $\lambda\ddot{\mathbf{x}} + 2\dot{\lambda}\dot{\mathbf{x}} + \ddot{\lambda}\mathbf{x}$

$+ \lambda K\mathbf{x} = 0$. Take the inner product with $\mathbf{x}$, remembering $(\mathbf{x}, \dot{\mathbf{x}}) = 0$ and $(\dot{\mathbf{x}}, \dot{\mathbf{x}}) + (\mathbf{x}, \ddot{\mathbf{x}}) = 0$. Thus follows $\ddot{\lambda} + \{(\mathbf{x}, K\mathbf{x}) - (\dot{\mathbf{x}}, \dot{\mathbf{x}})\}\lambda = 0$. Hence, $\phi(s) = (\mathbf{x}, K\mathbf{x}) - (\dot{\mathbf{x}}, \dot{\mathbf{x}})$ satisfies, and is continuous. In case $\mathbf{y} \not\equiv 0$, $\phi(s)$ is unique. Furthermore, $\phi \leqslant (\mathbf{x}, K\mathbf{x}) \leqslant m$.

LEMMA 3. *If $K(s)$ is as before; m as in Lemma 2, then $ml^2 < \pi^2$ implies that no non-trivial solution of $\ddot{\mathbf{y}} + K(s)\mathbf{y} = 0$, with $\mathbf{y}(0) = 0$ has a zero in the interval $0 < s \leqslant l$. Furthermore, for all $s \in [0, l]$, every solution with $\mathbf{y}(0) = 0$ satisfies*

$$|\mathbf{y}(s)| \leqslant \frac{\sin m^{\frac{1}{2}}s}{\sin m^{\frac{1}{2}}l}\,|\mathbf{y}(l)|, \qquad m > 0,$$

and

$$|\mathbf{y}(s)| \leqslant \frac{s}{l}\,|\mathbf{y}(l)|, \qquad m \leqslant 0.$$

Proof. There is no restriction in assuming $m \geqslant 0$ and $\mathbf{y} \not\equiv 0$. Every zero of $\mathbf{y}(s)$ is a zero of $|\mathbf{y}(s)| = \pm\lambda(s)$. λ satisfies $\ddot{\lambda} + \phi\lambda = 0$; $\lambda(0) = 0$, with $\phi \leqslant m$. The function $\psi(s) = \alpha \sin m^{\frac{1}{2}}s$ if $m > 0$, and $\psi(s) = \beta s$ if $m = 0$ (α and β any non-zero numbers) are solutions of $\ddot{\psi} + m\psi = 0$, with $\psi(0) = 0$. The smallest positive zero of the first function is $\pi m^{-\frac{1}{2}}$, which does not lie in $[0, l]$ because $\pi m^{-\frac{1}{2}} > (ml^2)^{\frac{1}{2}}m^{-\frac{1}{2}} = l$. Thus neither function ψ has a positive zero in the interval $[0, l]$. A well-known Sturmian theorem states that every positive zero of λ is preceded by one of ψ, which means, in this case, that λ has no positive zero in $[0, l]$. In order to prove this, *suppose* that λ has positive zeros in $[0, l]$; let s' be the smallest one. One may assume without loss of generality that $\lambda(s) \geqslant 0$ and $\psi(s) \geqslant 0$ for $s \in [0, s']$; then $\lambda(s') = 0$, $\dot{\lambda}(s') < 0$, $\psi(s') > 0$. Further, $m - \phi$ is strictly positive in some sub-interval of $[0, s']$; otherwise $m = \phi$ in $[0, s']$; and ψ is a multiple of λ; hence $\psi(s') = \lambda(s') = 0$, which is false. We have now

$$\dot{\lambda}\psi - \lambda\dot{\psi}\Big|_0^{s'} = \int_0^{s'} (\ddot{\lambda}\psi - \lambda\ddot{\psi})\,ds = \int_0^{s'} \lambda\psi(m - \phi)\,ds > 0,$$

and also

$$\begin{aligned}\dot{\lambda}\psi - \lambda\dot{\psi}\Big|_0^{s'} &= \dot{\lambda}(s')\psi(s') - \lambda(s')\dot{\psi}(s') \\ &\quad - \dot{\lambda}(0)\psi(0) + \lambda(0)\dot{\psi}(0) = \dot{\lambda}(s')\psi(s') < 0.\end{aligned}$$

Thus the assumption that λ has zeros in $[0, l]$ leads to a contradiction. This proves the first statement of the lemma; the second part uses the same formulas. The function λ/ψ is defined and of class C^1 for all $s \in [0, l]$; for $s = 0$ by limit procedure. Take $\lambda \geqslant 0$ in $[0, l]$, and

$$\psi(s) = \frac{\sin sm^{\frac{1}{2}}}{\sin lm^{\frac{1}{2}}}\lambda(l), \qquad \psi(s) = \frac{s}{l}\lambda(l)$$

respectively; then $\lambda(l) = \psi(l)$. For any s, with $0 < s \leqslant l$, we have

$$\psi^2\left(\frac{\lambda}{\psi}\right)'\Big|_0^s = \lambda\psi - \lambda\psi\Big|_0^s = \int_0^s \lambda\psi(m - \phi) > 0.$$

Since $\psi(0) = 0$, we have $\psi^2(0)(\lambda/\psi)'_0 = 0$; and thus follows $(\lambda/\psi)' > 0$ for all s, and λ/ψ is non-decreasing. Consequently,

$$\left(\frac{\lambda}{\psi}\right)_s < \left(\frac{\lambda}{\psi}\right)_l = 1;$$

that is, $\lambda(s) < \psi(s)$. This completes the second part.

LEMMA 4. *For every value of t in the interval $0 \leqslant t \leqslant 1$ the function ψ; defined by $\psi(t, \alpha) = \sin t\alpha/\sin\alpha$ when $\alpha \neq 0$, and by $\psi(t, 0) = t$ when $\alpha = 0$, is a monotone non-decreasing function of α in the interval $0 \leqslant \alpha < \pi$.*

Proof. The cases $t = 0$ and $t = 1$ are trivial. Take $0 < t < 1$; $0 < \alpha < \pi$, then

$$\frac{\partial\psi(t, \alpha)}{\partial\alpha} = \frac{N}{\sin^2\alpha},$$

where

$$N = t\cos t\alpha \sin\alpha - \sin t\alpha \cos\alpha = \sin(1 - t)\alpha - (1 - t)\cos t\alpha \sin\alpha$$

$$> \sin\alpha(1 - t) - (1 - t)\sin\alpha = \alpha(1 - t)\left[\frac{\sin(1 - t)\alpha}{(1 - t)\alpha} - \frac{\sin\alpha}{\alpha}\right] > 0,$$

because $\sin\alpha/\alpha$ is a decreasing function of α in the interval $0 \leqslant \alpha \leqslant \pi$. Hence, $\partial\psi/\partial\alpha > 0$ when $0 < \alpha < \pi$; since $\psi(t, \alpha)$ is continuous in α at $\alpha = 0$, it is monotone non-decreasing in α in the interval $0 \leqslant \alpha < \pi$.

3. Geodesic segments connecting a fixed point with the points of another curve. We return to the geometric situation of § 1. In this section, the results are true under the assumption $kD^2 < \pi^2$. Not until the proof in § 4 do we need the stronger condition $kD^2 < \pi^2/4$. When $k = 0$, $D = +\infty$, the expression kD^2 is interpreted as zero.

LEMMA 5. *In a convex set V in M, whose diameter and maximum curvature k satisfy $kD^2 < \pi^2$, let x be any point, and y_τ $(0 \leqslant \tau \leqslant 1)$ a C^1 path. The geodesic segment xy_τ is parametrized by t from 0 to 1 proportional to arc length: $(t, \tau) \to f(t, \tau)$; $f(0, \tau) = x$, $f(1, \tau) = y_\tau$. If*

$$v(t, \tau) = \frac{\partial f(t, \tau)}{\partial\tau}$$

denotes the tangent vector to the path $\tau \to f(t, \tau)$, for fixed t; and l_τ denotes the length of xy_τ, then

$$|v(t, \tau)| < \frac{\sin tl_\tau k^{1/2}}{\sin l_\tau k^{1/2}}\,|v(1, \tau)|$$

if $k > 0$; and

$$|v(t,\tau)| \leqslant t|v(1,\tau)| \quad \text{if } k \leqslant 0.$$

Proof. If s denotes the arc length function on xy_τ, measured from x to y_τ, and $\delta/\delta s$ denotes covariant differentiation with respect to s, then the components $v^\varkappa$ of v satisfy the classical Jacobi equation

$$\frac{\delta^2}{\delta s^2} v^\varkappa + v^\nu K^{\cdots\varkappa}_{\nu\mu\lambda} i^\mu i^\lambda = 0,$$

where i is the unit tangent vector to xy_τ. Choose along xy_τ a self-parallel orthonormal moving frame $(e_1, \ldots, e_n)$, with $i = e_1$. Denote the components of v with respect to this frame by $\mathbf{y} = (v^1, \ldots, v^n)$, and let $K(s)$ be the linear transformation

$$\mathbf{y} \to K(s)\mathbf{y} = (v^\nu K^{\cdots 1}_{\nu 11}, v^\nu K^{\cdots 2}_{\nu 11}, \ldots, v^\nu K^{\cdots n}_{\nu 11}),$$

where now all components of $K^{\cdots\varkappa}_{\nu\mu\lambda}$ are taken with respect to $(e_1, \ldots, e_n)$. Then $\mathbf{y} + K(s)\mathbf{y} = 0$; with, of course, $\mathbf{y}(0) = 0$. We claim that $m = \max\{k, 0\}$ is an upper bound for the inner product $(\mathbf{u}, K(s)\mathbf{u})$, with $|\mathbf{u}| = 1$. Decompose $\mathbf{u}$ into $\mathbf{u}'$ along i, and $\mathbf{u}''$ perpendicular to i. Then

$$\begin{aligned}(\mathbf{u}, K(s)\mathbf{u}) = K_{\varkappa\lambda\mu\nu} u^\varkappa i^\lambda i^\mu u^\nu &= -K_{\varkappa\lambda\mu\nu} u^\varkappa i^\lambda u^\mu i^\nu = K(\mathbf{u}, \mathbf{i}) \\ &= K(\mathbf{u}'', \mathbf{i}) \leqslant |\mathbf{u}''| \cdot k \leqslant |\mathbf{u}''| \cdot m \leqslant m.\end{aligned}$$

Since $l_\tau < D$, we have $ml_\tau^2 < \pi^2$, and thus, by Lemma 3, we have

$$|\mathbf{y}(s)| \leqslant \frac{\sin sm^{\frac{1}{2}}}{\sin l_\tau m^{\frac{1}{2}}} |\mathbf{y}(l_\tau)|, \qquad |\mathbf{y}(s)| \leqslant \frac{s}{l_\tau} |\mathbf{y}(l_\tau)|$$

respectively for $k = m > 0$ and $m = 0$ $(k \leqslant 0)$. In terms of the parameter $t = s/l_\tau$ this is precisely the statement of the lemma.

LEMMA 6. *Let $V, k, D, x, y_\tau, f(t,\tau)$ be as in Lemma 5, but let $y_0y_1\colon \tau \to f(1,\tau) = y_\tau$ be a geodesic segment of length L, parametrized by τ proportional to arc length. Then*

$$\rho(f(t,0), f(t,1)) \leqslant \frac{\sin tDk^{\frac{1}{2}}}{\sin Dk^{\frac{1}{2}}} L, \qquad \rho(f(t,0), f(t,1)) \leqslant tL$$

respectively in the cases $k > 0$ and $k \leqslant 0$.

Proof. We have $\rho(y_0, y_1) = L$, $|v(1,t)| = L$; and $\rho(f(t,0), f(t,1))$ is majorized by the length of the path $\tau \to f(t,\tau)$, $\tau \in [0,1]$. The latter equals $\int_0^1 |v(t,\tau)| d\tau$. In case $k = m > 0$, we have $l_\tau k^{\frac{1}{2}} < Dk^{\frac{1}{2}} < \pi$. Lemmas 4 and 5 (with $\alpha = l_\tau k^{\frac{1}{2}}$ and $\alpha = Dk^{\frac{1}{2}}$) are now applied:

$$\rho(f(t,0), f(t,1)) \leqslant \int_0^1 |v(t,\tau)| d\tau \leqslant L \int_0^1 \frac{\sin tl_\tau k^{\frac{1}{2}}}{\sin l_\tau k^{\frac{1}{2}}} d\tau \leqslant \frac{\sin tDk^{\frac{1}{2}}}{\sin Dk^{\frac{1}{2}}} L,$$

$$\rho(f(t,0), f(t,1)) \leqslant L \int_0^1 t d\tau = tL$$

(cases $k > 0$ and $k \leqslant 0$ respectively).

4. Proof of Theorem II. Choose a positive number ϵ. Let $y_1, y_2, z_1, z_2 \ni V$; and $\rho(y_1, y_2) < \epsilon$, $\rho(z_1, z_2) < \gamma\epsilon$, where $\gamma = \cos Dk^{\frac{1}{2}}$ if $k > 0$, and $\gamma = 1$ if $k \leqslant 0$. If $\phi, \chi, \psi: [0, 1] \to V$ are parametrizations proportional to arc length of the geodesic segments y_1z_1, y_2z_1, y_2z_2 respectively, then by Lemma 6 we have, for $k > 0$, $Dk^{\frac{1}{2}} < \frac{1}{2}\pi$:

$$\begin{aligned}
\rho(\phi(t), \psi(t)) &\leqslant \rho(\phi(t), \chi(t)) + \rho(\phi(t), \chi(t)) \\
&\leqslant \rho(y_1, y_2)\frac{\sin(1-t)Dk^{\frac{1}{2}}}{\sin Dk^{\frac{1}{2}}} + \rho(z_1, z_2)\frac{\sin tDk^{\frac{1}{2}}}{\sin Dk^{\frac{1}{2}}} \\
&< \frac{\epsilon \sin(1-t)Dk^{\frac{1}{2}} + \epsilon \cos Dk^{\frac{1}{2}} \sin tDk^{\frac{1}{2}}}{\sin Dk^{\frac{1}{2}}} \\
&= \epsilon\frac{\sin Dk^{\frac{1}{2}} \cos tDk^{\frac{1}{2}} - \cos Dk^{\frac{1}{2}} \sin tDk^{\frac{1}{2}} + \cos Dk^{\frac{1}{2}} \sin tDk^{\frac{1}{2}}}{\sin Dk^{\frac{1}{2}}} \\
&= \epsilon \cos tDk^{\frac{1}{2}} \leqslant \epsilon;
\end{aligned}$$

and for $k \leqslant 0$: $\rho(\phi(t), \psi(t)) < (1-t)\epsilon + t\epsilon = \epsilon$. This completes the proof.

Final Remark. The lemmas derived in §§ 2 and 3, can be used to deduce more results on upper bounds for the distance between corresponding points of geodesic segments in convex sets with either $kD^2 < \pi^2$ or $kD^2 < \frac{1}{4}\pi^2$. The approach to take seems so obvious that there is no need here to amplify this point with a number of examples.

References

1. J. H. C. Whitehead, *Convex regions in the geometry of paths*, Quart. J. Math., Oxford, Ser. *3* (1932), 33–42.
2. E. A. Michael, *Convex structures and continuous selections*. Can. J. Math., *11* (1959), 556–575.
3. H. Busemann, *The geometry of geodesics* (New York), 1955.

University of Washington

THE a-LOCAL OPERATOR PROBLEM

LOUIS DE BRANGES

Let $f(t) = \int e^{itx} d\mu(x)$ be the Fourier transform of a Borel measure of finite total variation. The formula

$$f'(t) = \int e^{itx} ix \, d\mu(x)$$

can be justified if the integral on the right converges absolutely. For

$$\frac{f(t+h) - f(t)}{h} = \int e^{itx} \frac{e^{ixh} - 1}{ixh} ix \, d\mu(x)$$

where

$$\left| \frac{e^{ixh} - 1}{ixh} \right| = \left| \frac{2}{xh} \sin \frac{xh}{2} \right| < 1.$$

Now let $h \to 0$ in both sides of this equation and use the Lebesgue dominated convergence theorem.

The formula suggests a concept of "derivative" for such absolutely convergent Fourier transforms. Notationally it is easier to drop a factor of i. Define an operator H to act on $f(t)$ by

$$H \cdot f(t) = \int e^{itx} x \, d\mu(x)$$

whenever

$$\int |x \, d\mu(x)| < \infty.$$

The operator H corresponds to $-i$ times differentiation and we have the formula

$$H \cdot f(t) = \lim_{h \to 0} \frac{f(t+h) - f(t)}{ih}$$

whenever $f(t)$ is in the domain of H. In **(2)**, this formula was shown to have interesting consequences from the point of view of the operational calculus.

Let $K(x)$ be a complex valued, Borel measurable function of real x. Define a corresponding operator $K(H)$ on Fourier transforms

$$f(t) = \int e^{itx} d\mu(x)$$

of Borel measures of finite total variation by

$$K(H) \cdot f(t) = \int e^{itx} K(x) d\mu(x)$$

whenever $\int |K(x) d\mu(x)| < \infty$.

Received July 11, 1958. This work was started at Cornell University under the supervision of H. Pollard.

For example, if $K(H) = H^2$,

$$\begin{aligned} H^2 \cdot f(t) &= \int e^{ixt} x^2 \, d\mu(x) \\ &= \int e^{ixt} x (x \, d\mu(x)) \\ &= H \cdot Hf(t) \end{aligned}$$

and H^2 is the double iteration of H. Similarly, if $K(H)$ is any polynomial in H, the operator $K(H)$ may be thought of as a finite linear combination of iterations of H. The functional notation $K(H)$ which is so appropriate in these simple cases is also to be used when $K(x)$ is only restricted to be Borel measurable.

Let $a > 0$. The operator $K(H)$ is said to be a-local if whenever $f(t)$ in the domain of $K(H)$ vanishes in $[-a, a]$,

$$K(H) \cdot f(0) = 0.$$

The problem is to find conditions on the defining function $K(x)$ that the associated operator $K(H)$ be a-local.

The problem here is related to the problem in **(2)** of finding local operators on Fourier transforms, and was proposed by Pollard at the same time as the other problem. We continue with the ideas and notation of the previous paper, except that Fourier transforms of not absolutely continuous measures are admitted in the domain of the operator. The defining function $K(x)$ is assumed to have a finite value for each real x, and there are no a.e. identifications.

Recall the lemma of **(2)** which is the basis of our approach.

LEMMA 1. *If $K(z)$ is an entire function of exponential type such that*

$$\overline{\lim_r} \, r^{-1} \log |K(re^{i\theta})| \leqslant a|\sin\theta|, \tag{1}$$

and

$$e^{-a|y|}K(iy) \qquad (z = x + iy) \tag{2}$$

is bounded, and if

$$f(t) = \int e^{ixt} d\mu(x)$$

is an absolutely convergent Fourier transform in the domain of $K(H)$ which vanishes in $[-a, a]$, then for all complex z

$$\int \frac{K(t) - K(z)}{t - z} \, d\mu(t) = 0.$$

Theorems I and II are the appropriate restatements of Theorem I of **(2)**.

THEOREM I. *If $K(z)$ is an entire function of exponential type satisfying* (1) *and* (2), *then the operator $K(H)$ is a-local.*

THEOREM II. *If $K(H)$ is an a-local operator on Fourier transforms and if $K(H)$ has in its domain a function which vanishes in $[-a, a]$ and does not*

vanish identically, then $K(x)$ is the restriction to the real axis of an entire function of exponential type satisfying (1) *and*

$$\int \frac{\log^+|K(t)|}{1+t^2}\,dt < \infty. \tag{3}$$

The gap between Theorem I and Theorem II arising from (2) has caused difficulty. The hypothesis (2) cannot be entirely removed from Theorem I. In fact, the function $K(z) = z \sin az$ does not define an a-local operator on Fourier transforms. This follows from the following uniqueness theorem.

THEOREM III. *If $K(H)$ is a-local and if*

$$\sum_{-\infty}^{+\infty} (1+n^2)^{-1}|K(n\pi/a)| < \infty,$$

then $K(x)$ is the restriction to the real axis of an entire function $K(z)$ of exponential type satisfying (1) *and*

$$e^{-a|y|}K(iy) = o(y) \tag{4}$$

as $|y| \to \infty$.

Although Theorem III shows that the hypothesis (2) cannot be entirely removed from Theorem I, we conjecture that it can be removed if some hypothesis is made which keeps the modulus of $K(z)$ from being too bumpy. We have been able to obtain the conclusion of Lemma 1 from a hypothesis of this nature.

LEMMA 2. *If $K(z)$ is an entire function of exponential type satisfying* (1) *and* (3) *and*

$$\frac{\partial}{\partial y}\frac{y}{\pi}\int \frac{\log^+|K(t)|}{(t-x)^2+y^2}\,dt < c \tag{5}$$

is bounded above in the upper half plane $y > 0$, and if

$$f(t) = \int e^{ixt}\,d\mu(x)$$

in the domain of $K(H)$ vanishes in $[-a, a]$, then for all complex z

$$\int \frac{K(t) - K(z)}{t - z}\,d\mu(t) = 0.$$

We have not been able to show that under the hypotheses of Lemma 2, the operator $K(H)$ is a-local. The trouble lies in the following property of a-local operators which we have been unable to relate to any explicit description of $K(z)$.

THEOREM IV. *If $K(z)$ is an entire function of exponential type such that*

$$\overline{\lim_r}\, r^{-1}\log K(re^{i\theta}) = a|\sin\theta|, \tag{6}$$

and if the operator $K(H)$ is a-local, and if μ is a Borel measure on the real line, of finite total variation, such that

$$\int |K(t)d\mu(t)| < \infty$$

and for all complex z

$$\int \frac{K(t) - K(z)}{t - z}\, d\mu(t) = 0,$$

then

$$\int K(t)d\mu(t) = 0.$$

It is interesting to compare the hypotheses of Lemma 2 with the following alternative hypotheses for Levinson's theorem (Theorem II of **(2)**).

THEOREM V. *If* $K(x) \geq 1$ *is a continuous function of real* x *such that* $\log K(x)$ *is uniformly continuous and*

$$\int \frac{\log K(x)}{1 + x^2}\, dx = \infty, \tag{7}$$

then there is no non-zero Borel measure μ *such that*

$$\int |K(x)d\mu(x)| < \infty, \quad \text{and} \quad \int e^{itx} d\mu(x)$$

vanishes in an interval.

Proof of Theorem I. Let $f(t) = \int e^{itx} d\mu(x)$ be in the domain of $K(H)$ and vanish in $[-a, a]$. By Lemma 1,

$$\begin{aligned}\int K(t)d\mu(t) &= \int \frac{tK(t) - zK(t)}{t - z}\, d\mu(t) \\ &= \int \frac{tK(t) - zK(z)}{t - z}\, d\mu(t) \\ &= \int \frac{tK(t) - ze^{ias}K(z)e^{-ias}}{t - z}\, d\mu(t).\end{aligned}$$

Since

$$0 = \int e^{-ias} d\mu(t) = \int \frac{te^{-ias} - ze^{-ias}}{t - z}\, d\mu(t),$$

$$\int K(t)d\mu(t) = \int \frac{e^{ias}K(t) - e^{ias}K(z)}{t - z}\, te^{-ias} d\mu(t).$$

Let $z = iy$ where $y \to +\infty$. By the Lebesgue dominated convergence theorem,

$$\int (t - z)^{-1} e^{ias} K(t) te^{-ias} d\mu(t) \to 0$$

$$\int (t - z)^{-1} te^{-ias} d\mu(t) \to 0$$

and since by (2) $e^{ias}K(z)$ remains bounded in the limit,

$$K(H) \cdot f(0) = \int K(t)d\mu(t) = 0. \qquad \text{q.e.d.}$$

Proof of Theorem II. By hypothesis there is a function

$$f_1(t) = \int e^{itx} d\mu_1(x)$$

in the domain of $K(H)$ which vanishes in $[-a, a]$ and does not vanish identically. Write

$$f_1(t) = \int e^{itx} B_1(x) d\mu(x)$$

where μ is a non-negative Borel measure and $B_1(x)$ is a Borel measurable function of absolute value 1. The hypothesis that $f_1(t)$ vanishes in $[-a, a]$ implies that the closed span in $L^1(\mu)$ of the functions e^{itx}, $-a \leqslant t \leqslant a$, is not all of $L^1(\mu)$. Since $f_1(t)$ is in the domain of $K(H)$, $K(x)$ is in $L^1(\mu)$. The hypothesis that $K(H)$ is a-local implies that for every Borel measurable function $B(x)$ which is essentially bounded with respect to μ, such that $\int e^{itx} B(x) d\mu(x) = 0$ for $-a \leqslant t \leqslant a$, we have $\int K(x) B(x) d\mu(x) = 0$. Since every continuous linear functional on $L^1(\mu)$ is defined by a Borel measurable function which is essentially bounded with respect to μ, it follows by the Hahn-Banach theorem that $K(x)$ lies in the closed span in $L^1(\mu)$ of the functions e^{itx}, $-a \leqslant t \leqslant a$. Let

$$M(z) = \sup |L(z)|$$

where $L(z)$ ranges in the finite linear combinations of the functions e^{itz} where $-a \leqslant t \leqslant a$. Since $L(z) = (\int d\mu)^{-1}$ appears in the supremum, $M(z) \geqslant (\int d\mu)^{-1}$. Since the closed span of the functions e^{itx}, $-a \leqslant t \leqslant a$, is not all of $L^1(\mu)$, it follows as in the proof of necessity of Theorem I of **(2)** that

$$\int \frac{\log M(t)}{1+t^2} dt < \infty \tag{8}$$

and

$$\log M(x+iy) \leqslant a|y| + \frac{y}{\pi} \int \frac{\log M(t)}{(t-x)^2 + y^2} dt \qquad (y \neq 0) \tag{9}$$

and that $K(x)$ is equal a.e. with respect to μ to the restriction to the real axis of a unique entire function $F(z)$ such that for all complex z

$$F(z) \leqslant M(z) \int |K(t)| d\mu(t). \tag{10}$$

Let μ' be any non-negative Borel measure of finite total variation such that $\mu \leqslant \mu'$ and $\int |K(t)| d\mu'(t) < \infty$, and let $M'(z)$ be the corresponding majorant for μ'. Since for any Borel function $L(t)$,

$$\int |L(t)| d\mu(t) \leqslant \int |L(t)| d\mu'(t),$$

we have for all complex z,

$$M'(z) \leqslant M(z).$$

Therefore, the closed span in $L^1(\mu')$ of the functions, e^{itz} $-a \leqslant t \leqslant a$, is not all of $L^1(\mu')$ and $K(x)$ is equal a.e. with respect to μ' to the restriction to the real axis of the entire function $F(z)$, and for all complex z

$$F(z) \leqslant M'(z) \int |K(t)| d\mu'(t).$$

By the arbitrariness in the choice of μ', $K(x) = F(x)$ for all real x. (1) and (3) now follow from (8), (9), and (10), q.e.d.

Proof of Theorem III. Let x_1 and x_2 be any two distinct real numbers, neither of which is an integral multiple of π/a. Let μ be the measure which has mass $(\sin ax_2)^{-1}$ at x_2, mass $-(\sin ax_1)^{-1}$ at x_1, and for integral n has mass

$$\frac{(-1)^n (ax_2 - ax_1)}{(ax_2 - n\pi)(ax_1 - n\pi)}$$

at $n\pi/a$. It is easy to verify that $\int |d\mu(x)| < \infty$ and $\int e^{itx} d\mu(x) = 0$ for $-a \leqslant t \leqslant a$. (Use the Fourier series for e^{itx} in $-a \leqslant t \leqslant a$:

$$e^{itx} \sim \Sigma (-1)^n (ax - n\pi)^{-1} \sin ax\, e^{in\pi t/a}.)$$

Let $K(x)$ satisfy the hypotheses of the theorem. They imply that $\int |K(t) d\mu(t)| < \infty$. Since the operator $K(H)$ is a-local, $\int K(x) d\mu(x) = 0$. Equivalently,

$$\frac{K(x_2)}{\sin ax_2} - \frac{K(x_1)}{\sin ax_1} = \sum \frac{(-1)^n (ax_1 - ax_2) K(n\pi/a)}{(ax_2 - n\pi)(ax_1 - n\pi)}.$$

Define $K(z)$ for complex z in the unique way such that the same formula holds with x_1 and x_2 replaced by complex variables z_1 and z_2. Obvious estimates from this formula show that $K(z)$ is an entire function of exponential type satisfying (1) and (4). (A similar formula is discussed in **(1**, pp. 220–1.**)**

Proof of Lemma 2. Let $f(t)$ satisfy the hypotheses of the lemma and let

$$L(z) = \int \frac{K(t) - K(z)}{t - z}\, d\mu(t). \tag{11}$$

As in the proof in **(2)** of Lemma 1, $L(z)$ is an entire function of minimal exponential type. Define

$$\log S(x) = \log^+ |K(x)|$$

$$\log S(x + iy) = \frac{|y|}{\pi} \int \frac{\log S(t)}{(t - x)^2 + y^2}\, dt \qquad (y \neq 0)$$

The convergence follows from (3) and by (5)

$$S(x + iy) \leqslant e^{a|y|} S(x). \tag{12}$$

Let the real number t be held fixed and consider

$$\frac{K(t) - K(z)}{t - z}$$

as a function of complex z. Then

$$\left| \frac{K(t) - K(z)}{t - z} \right| \leqslant |y|^{-1} (|K(t)| + |K(z)|)$$

where by Boas **(1**, p. 93**)**,

$$|K(z)| \leqslant S(z) e^{a|y|}.$$

Since $S(t)$, $S(z) \geqslant 1$, we have

$$\left|\frac{K(t) - K(z)}{t - z}\right| < 2|y|^{-1} S(t)S(z)e^{a|y|}.$$

Since

$$\frac{K(t) - K(z)}{t - z}$$

has the same growth properties as $K(z)$, the same reference to Boas shows that

$$\begin{aligned}\left|\frac{K(t) - K(x)}{t - x}\right| &< 2|y|^{-1} S(t)S(x + 2iy)e^{2a|y|} \\ &< 2|y|^{-1} e^{3(c+a)|y|} S(t)S(x)\end{aligned}$$

by (12). Choosing $y = (2c + 2a)^{-1}$, we have

$$\left|\frac{K(t) - K(x)}{t - x}\right| < 4e(c + a)S(t)S(x) \tag{13}$$

and hence by (11)

$$|L(x)| < 4e(c + a)S(x)||S\mu|| \tag{14}$$

where

$$||S\mu|| = \int S(t)|d\mu(t)| < \infty.$$

We claim that for $n = 1, 2, 3, \ldots$

$$|L(x)|^n < [4e(c + a)||S\mu||]^n S(x) \tag{15}$$

and hence (by the Boas reference)

$$|L(z)|^n < [4e(c + a)||S\mu||]^n S(z). \tag{16}$$

We prove the inequality (15) by induction on n, starting with (14) when $n = 1$. Suppose that for some n, (15) holds, and we will prove the corresponding inequality with n replaced by $n + 1$. Since $L(z)$ has minimal exponential type, so has $L^n(z)$, and by (15), $\int |L^n(t)d\mu(t)| < \infty$. By Lemma 1,

$$\int \frac{L^n(t) - L^n(z)}{t - z}\, d\mu(t) = 0$$

for all complex z. Therefore,

$$\begin{aligned}L^{n+1}(z) &= \int \frac{L^n(z)K(t) - K(z)L^n(z)}{t - z}\, d\mu(t) \\ &= \int \frac{L^n(z)K(t) - K(z)L^n(t)}{t - z}\, d\mu(t).\end{aligned}$$

By (16),

$$\begin{aligned}|L^{n+1}(z)| &< |y|^{-1}[|L^n(z)|\textstyle\int |K(t)d\mu(t)| + |K(z)|\int |L^n(t)d\mu(t)|] \\ &< 2|y|^{-1}[4e(c + a)||S\mu||]^n S(z)e^{a|y|}||S\mu||.\end{aligned}$$

The conclusion

$$L^{n+1}(x) \leqslant [4e(c+a)||S\mu||]^{n+1}S(x)$$

now follows in the same way we established (13). This completes the inductive step.

Since in (16), n is arbitrary, for all complex z,

$$|L(z)| \leqslant 4e(c+a)||S\mu||.$$

By Liouville's theorem, $L(z) = L$ is a constant.

Consider various cases depending on the number and location of the zeros of $K(z)$. Discard the case that $K(z)$ has only a finite number of zeros, for it follows by Lemma 1. In this case, $K(z) = P(z)e^{ihz}$ for some polynomial $P(z)$ and $-a \leqslant h \leqslant a$. If $|h| < a$, Lemma 1 can be used directly. For the limiting cases, let $h \to a$ or $-a$ and use dominated convergence.

If $K(z)$ admits an infinite sequence of zeros $z_n = x_n + iy_n$ such that $|z_n| \to \infty$ and y_n is bounded away from 0, then

$$L = \lim_n \int \frac{K(t)}{t - z_n}\, d\mu(t) = 0$$

by dominated convergence.

We are left with the case that $K(z)$ admits an infinite set of zeros, but all so close to the real axis that the last argument does not apply. Let $h > 0$. By (12), $K(t + ih)$ is dominated on the real axis by $S(t)$ and the above argument with $K(z)$ replaced by $K(z + ih)$ shows that

$$\int \frac{K(t+ih) - K(z+ih)}{t - z}\, d\mu(t)$$

is a constant, independent of z. Since $K(z + ih)$ has an infinite set of zeros at a positive distance from the real axis, the constant is 0. On letting $h \to 0$ and using (12), we have by the Lebesgue dominated convergence theorem,

$$\int \frac{K(t) - K(z)}{t - z}\, d\mu(t) = 0.$$

Proof of Theorem IV. Let K and μ be as in the hypotheses of the theorem, and let $0 < h < a$ be held fixed. As in the proof of Lemma 1,

$$L(z) = \int \frac{e^{ith} - e^{itz}}{t - z}\, d\mu(t)$$

is an entire function of exponential type. By the hypotheses on $K(z)$,

$$\begin{aligned} K(z)L(z) &= \int \frac{K(z)e^{ith} - e^{itz}K(z)}{t - z}\, d\mu(t) \\ &= \int \frac{K(z)e^{ith} - e^{itz}K(t)}{t - z}\, d\mu(t) \end{aligned}$$

and hence

$$|K(z)L(z)| \leqslant |y|^{-1}(|K(z)| + e^{-hy})||\mu||$$

where $||\mu|| = \int |d\mu| < \infty$. The hypothesis (6) now implies that

$$\overline{\lim_r}\, r^{-1} \log|K(re^{i\theta})L(re^{i\theta})| \leqslant a|\sin\theta|.$$

By applying the theorem **(1**, p. 116**)** on the effective asymptotic behaviour of these functions and using (6), we see that $L(z)$ has minimal exponential type. The theorem is applicable since $L(z)$ is bounded on a line

$$|L(z)| \leqslant 2|y|^{-1}||\mu||$$

and $K(z)$ is the ratio of two functions bounded and analytic in a half plane. By **(1**, p. 83**)**, the entire function $L(z)$, being of minimal exponential type and bounded on a line, is a constant. By dominated convergence

$$L(z) = L = \lim_{y\to+\infty} \int \frac{e^{ixt} - e^{-hy}}{t - iy}\, d\mu(t) = 0.$$

Again by dominated convergence,

$$\begin{aligned}\int e^{ixt} d\mu(t) &= \lim_{y\to+\infty} -iy \int \frac{e^{ixt}}{t - iy}\, d\mu(t) \\ &= \lim_{y\to+\infty} -iy \int \frac{e^{-hy}}{t - iy}\, d\mu(t) \\ &= 0.\end{aligned}$$

A similar argument shows that

$$f(h) = \int e^{ixt} d\mu(t)$$

vanishes when $-a < h < 0$. By the continuity of the Fourier transforms, this function vanishes in $-a \leqslant h \leqslant a$. Since the operator $K(H)$ is a-local,

$$\int K(t)\, d\mu(t) = K(H)\cdot f(0) = 0.$$

Proof of Theorem V. Let $a > 0$ and let $T(x) = T_a(x) = \sup |L(x)|$ where $L(z)$ ranges in the entire functions of exponential type, satisfying (1) and (2), and such that for all real x, $|L(x)| \leqslant K(x)$. By the proof of Theorem II of **(2)**, the conclusion of the theorem follows if we can show that

$$\int \frac{\log T(x)}{1 + x^2}\, dx = \infty \tag{17}$$

(for every $a > 0$). By the uniform continuity of $\log K(x)$, there is some $\epsilon > 0$ such that

$$|\log K(x_1) - \log K(x_2)| \leqslant \epsilon$$

whenever $|x_2 - x_1| \leqslant \pi/2a$. Let the real number x_0 be held fixed. For every positive integer n,

$$\log K(x) \geqslant \log K(x_0) - n\epsilon$$

whenever $|x - x_0| \leqslant n\pi/2a$, and we will use this estimate with n so chosen that

$$n\,(\log \pi/2 + \epsilon) < \log K(x_0)$$

(if such an n exists). For then the function

$$L(x) = \left[\frac{\pi n}{2a(x - x_0)} \sin \frac{a(x - x_0)}{n}\right]^n$$

is an entire function of exponential type, satisfying (1) and (2), and when $|x - x_0| < n\pi/2a$,

$$\log |L(x)| < n \log \pi/2 < \log K(x)$$

and when $|x - x_0| > n\pi/2a$,

$$\log |L(x)| < 0 < \log K(x).$$

So $L(x)$ is one of the test functions in the definition of $T(x)$. Since

$$\log L(x_0) = n \log \pi/2,$$

we have

$$\log T(x_0) > n \log \pi/2$$

whenever

$$n(\log \pi/2 + \epsilon) < \log K(x_0).$$

So, for all real x,

$$\log T(x) > \frac{\log \pi/2}{\log \pi/2 + \epsilon} \log K(x) - \log \pi/2.$$

The hypothesis (7) now implies (17).

References

1. R. P. Boas, Jr., *Entire Functions* (New York, 1954).
2. L. de Branges, *Local Operators on Fourier Transforms*, Duke Math. J. **25** (1958), 143–54.

Lafayette College

ON A THEOREM OF HAYMAN CONCERNING QUASI-BOUNDED FUNCTIONS

P. B. KENNEDY

1. If $f(z)$ is regular in $|z| < 1$, the expression

$$m(r, f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log^+ |f(re^{i\theta})| \, d\theta \qquad (0 < r < 1)$$

is called the *characteristic* of $f(z)$. This is the notation of Nevanlinna **(4)** for the special case of regular functions; in this note it will not be necessary to discuss meromorphic functions. If $m(r, f)$ is bounded for $0 < r < 1$, then $f(z)$ is called *quasi-bounded* in $|z| < 1$. In particular, every bounded function is quasi-bounded. The class Q of quasi-bounded functions is important because, for instance, a "Fatou theorem" holds for such functions **(4**, p. 134**)**. It is known that $f \in Q$ does not imply $f' \in Q$ in general; the simplest counterexample seems to be that given in **(1)**, §8. Hayman **(1**, Theorem IV**)** found extra hypotheses under which $f \in Q$ implies $f' \in Q$. The purpose of this note is to construct examples which shed some light on the "goodness" of these hypotheses. The method is similar to one used in **(2)**; it is essentially a combination of a construction due to Littlewood **(3**, §8.5**)** with a theorem of Specht **(5)** on conformal mapping.

2. Let D be a domain containing a sequence of open arcs

$$z = e^{i\theta}, \quad \theta_n < \theta < \theta_n + \delta_n, \quad n = 1, 2, 3, \ldots. \tag{2.1}$$

Let $d(\theta)$ denote the distance from $e^{i\theta}$ to the boundary of D, and suppose that there are positive numbers A and B, independent of θ and n, such that

$$d(\theta) > A\{(\theta - \theta_n)(\theta_n + \delta_n - \theta)\}^B \quad (\theta_n < \theta < \theta_n + \delta_n, \, n = 1, 2, 3, \ldots).$$

Then, following Hayman **(1)**, we say that D *properly contains* the sequence of arcs (2.1). Hayman has shown that, *if D contains $|z| < 1$ and properly contains a sequence of non-overlapping arcs* (2.1) *such that*

$$\sum_1^\infty \delta_n = 2\pi, \tag{2.2}$$

$$\sum_1^\infty \delta_n \log \frac{1}{\delta_n} < \infty, \tag{2.3}$$

Received October 14, 1958. Research supported in part by the Air Research and Development Command, United States Air Force, through its European Office, under Contract No. AF 61 (514) – 1399.

and if $f(z)$ is regular and bounded in D, then $f'(z)$ is quasi-bounded in $|z| < 1$. This is a very special case of Theorem IV of **(1)**, but it is sufficient for our purpose in this note.

We prove the following theorem, which sets a limit to the extent to which it may be possible to relax the condition (2.3).

THEOREM. *Let $\{\delta_n\}$ be a non-increasing sequence satisfying* (2.2) *and*

$$\limsup_{n\to\infty}\left(\sum_{j=n}^{\infty}\delta_j\right)\log\frac{1}{\delta_n}=\infty, \tag{2.4}$$

and let $\{\theta_n\}$ be such that no two of the arcs (2.1) *overlap. Then there exist a domain D containing $|z| < 1$ and properly containing the sequence of arcs* (2.1), *and a function $f(z)$ regular and bounded in D such that $m(r, f') \to \infty$ as $r \to 1$.*

(2.4) does *not* follow from

$$\sum_{1}^{\infty}\delta_n\log\frac{1}{\delta_n}=\infty; \tag{2.5}$$

for example, if

$$\delta_n=\frac{1}{n(\log n)^2}$$

for all large enough n, then (2.4) is false although (2.5) is true. Thus there remains a gap between the positive information given by Hayman's theorem and the negative information which comes from the result proved in this note. I have not found a method of closing this gap. However, our result shows that (2.3) cannot in general be replaced by

$$\sum_{1}^{\infty}\delta_n\left(\log\frac{1}{\delta_n}\right)^{\alpha}<\infty$$

for any $\alpha < 1$, or even by certain rather stronger conditions than this.

In the theorem, the sequence of arcs (2.1) may be obtained by putting $\theta_1 = 0$, $\theta_{n+1} = \theta_n + \delta_n$, in which case (2.1) is a sequence of arcs whose endpoints have no limit-point except $z = 1$, and whose complement is the sequence $\{e^{i\theta_n}\}$. But it should be noted that a more complicated situation may also occur, in which the arcs (2.1) form the complement of an uncountable set.

Some remarks on the other hypotheses of Hayman's theorem are made in §6.

3. In this section and the next we assume only that (2.1) is a sequence of non-overlapping arcs; the other hypotheses of the theorem are not needed at this stage. However, we suppose, as we may without loss of generality, that $\theta_1 \leqslant \theta_n < \theta_1 + 2\pi$ for every n. Let $c > 0$ be a constant, and for $n = 1, 2, 3, \ldots$ put

$$\psi(\theta)=c(\theta-\theta_n)^2(\theta_n+\delta_n-\theta)^2 \qquad (\theta_n<\theta<\theta_n+\delta_n).$$

Moreover put $\psi(\theta) = 0$ for all other values of θ in $\theta_1 \leqslant \theta < \theta_1 + 2\pi$, and define $\psi(\theta)$ outside this interval by requiring it to have period 2π. Then we have

LEMMA 1. *There is an absolute constant K with the following properties:*
(i) *for all θ and n,*

$$\psi(\theta) \leqslant Kc|\theta - \theta_n|^3,$$

(ii) *$\psi(\theta)$ is twice continuously differentiable, and for all θ,*

$$|\psi'(\theta)| \leqslant Kc, \qquad |\psi''(\theta)| \leqslant Kc.$$

A full proof of Lemma 1 involves some lengthy repetition, and so we leave some details to the reader. Let the subset S of $(\theta_1, \theta_1 + 2\pi)$ be defined by

$$S = \bigcup_1^\infty (\theta_n, \theta_n + \delta_n),$$

and let T denote the complement of S with respect to the closed interval $(\theta_1, \theta_1 + 2\pi)$. We first prove that, for $\phi \in T$ and all θ,

$$\psi(\theta) \leqslant K_1 c|\theta - \phi|^3, \tag{3.1}$$

where K_1 is an absolute constant. If $\theta \in S$, then $\theta_j < \theta < \theta_j + \delta_j$ for some j, and therefore, since $\phi \in T$,

$$|\theta - \phi| > \min\,(\theta - \theta_j, \theta_j + \delta_j - \theta).$$

Since

$$\theta - \theta_j < 2\pi, \quad \theta_j + \delta_j - \theta < 2\pi, \tag{3.2}$$

it follows that

$$|\theta - \phi| > \frac{1}{2\pi}(\theta - \theta_j)(\theta_j + \delta_j - \theta).$$

From this and the definition of $\psi(\theta)$ we get (3.1), with a suitable K_1, for all $\theta \in S$; and (3.1) is obvious if $\theta \in T$. Further, a simple argument from the periodicity of $\psi(\theta)$ shows that (3.1) must then be true for all θ. In particular, putting $\phi = \theta_n$, we get (i) with $K = K_1$.

From (3.1) it follows that $\psi'(\phi) = 0$ for all $\phi \in T$. For $\theta_n < \theta < \theta_n + \delta_n$ we have

$$\psi'(\theta) = 3c(\theta - \theta_n)^2(\theta_n + \delta_n - \theta)^2(2\theta_n + \delta_n - 2\theta).$$

We can now carry out an argument similar to that used above, with $\psi'(\theta)$ instead of $\psi(\theta)$, to find that

$$|\psi'(\theta)| \leqslant K_2 c|\theta - \phi|^2 \quad (\phi \in T, \text{ all } \theta),$$

where K_2 is an absolute constant. It follows easily that $\psi'(\theta)$ is everywhere continuous and satisfies $|\psi'(\theta)| \leqslant 4\pi^2 K_2 c$. Thus the part of (ii) which refers to $\psi'(\theta)$ is proved with $K = 4\pi^2 K_2$.

The required properties of $\psi''(\theta)$ are proved similarly. Lemma 1 then follows, for some sufficiently large absolute constant K.

Let the domain D be the interior of the simple closed contour $r = \exp \psi(\theta)$ in the $z = re^{i\theta}$-plane. Then we have the following lemma.

LEMMA 2. *D contains $|z| < 1$ and properly contains the sequence of arcs* (2.1).

It is obvious that D contains $|z| < 1$. Let $d(\theta, h)$ denote the distance between $e^{i\theta}$ and $\exp\{\psi(\theta + h) + i(\theta + h)\}$. To complete the proof of Lemma 2 it is plainly enough to show that there is a number $A > 0$, independent of θ, h, and n, such that

$$d(\theta, h) > A(\theta - \theta_n)^3(\theta_n + \delta_n - \theta)^3 \tag{3.3}$$

for all θ and h satisfying

$$\theta_n < \theta < \theta_n + \delta_n, \quad \theta_n < \theta + h < \theta_n + \delta_n. \tag{3.4}$$

By periodicity we may suppose

$$|h| \leqslant \pi. \tag{3.5}$$

When (3.4) is true, we have

$$\{d(\theta, h)\}^2 = \{\exp \psi(\theta + h) - 1\}^2 + 4 \sin^2 \tfrac{1}{2} h \exp \psi(\theta + h).$$

Thus by (3.5),

$$d(\theta, h) > 2|\sin \tfrac{1}{2} h| > 2\pi^{-1}|h|,$$

and if

$$|h| \geqslant (4\pi)^{-1}(\theta - \theta_n)(\theta_n + \delta_n - \theta), \tag{3.6}$$

then (3.3) follows from this with a suitable absolute constant A, say $A = A_0$.

On the other hand, if (3.6) is false, then by (3.2)

$$|h| < \tfrac{1}{2}(\theta - \theta_n), \quad |h| < \tfrac{1}{2}(\theta_n + \delta_n - \theta).$$

Therefore

$$\begin{aligned} d(\theta, h) &\geqslant \exp \psi(\theta + h) - 1 > \psi(\theta + h) \\ &> c(\theta - \theta_n - |h|)^3(\theta_n + \delta_n - \theta - |h|)^3 \\ &> 2^{-6}c(\theta - \theta_n)^3(\theta_n + \delta_n - \theta)^3, \end{aligned}$$

and this is (3.3) with $A = 2^{-6}c$. Thus in any case (3.3) is true with $A = \min(A_0, 2^{-6}c)$, and this proves Lemma 2.

4. We continue in the notation of §3. We now suppose that the constant c, occurring in the definition of $\psi(\theta)$, is chosen so small that

$$1 \leqslant \exp \psi(\theta) < \frac{81}{80}, \tag{4.1}$$

$$|\psi'(\theta)| < \frac{1}{80}, \tag{4.2}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\psi'(\theta)-\psi'(\theta_0)}{\sin\frac{1}{2}(\theta-\theta_0)}\right|d\theta<\frac{1}{80}, \tag{4.3}$$

$$\exp\psi(\theta)<2-r \quad \text{for } |\theta-\theta_n|<1-r \qquad (n=1,2,3,\dots), \tag{4.4}$$

the integral in (4.3) being uniformly convergent with respect to θ_0 for $0\leqslant\theta_0\leqslant 2\pi$. The possibility of such a choice of c is easily seen from Lemma 1.

Let $w=w(z)$ map D one-one and conformally onto $|w|<1$, with $w(0)=0$ and $w'(0)>0$. We then have the following lemma, due to Specht (**5**, Theorem III).

LEMMA 3. (4.1), (4.2), *and* (4.3) *together imply that* $w'(z)$ *exists* for all* z *in the closure of* D *and satisfies*

$$|w'(z)-1|<\tfrac{1}{2}. \tag{4.5}$$

Lemma 3 follows at once from Specht's result if, in that result, we interchange the z- and w-planes, write $\exp\psi(\theta)$ instead of $\rho(\theta)$, and put $\epsilon=\delta=1/80$. It should be noted that Specht actually gives a bound for $|z'(w)-1|$, where $z(w)$ is the inverse function of $w(z)$; but from this one may readily derive (4.5), which is more suitable for our purpose.

From Lemma 3 and (4.4) we now deduce

LEMMA 4. *For* $|\theta-\theta_n|<1-r$ *we have*

$$\tfrac{1}{2}(1-r)<1-|w(re^{i\theta})|<3(1-r).$$

First let $z(w)$ be the inverse function of $w(z)$, as above. If $re^{i\theta}=z(\rho e^{i\varphi})$, then

$$1-r\leqslant|z(e^{i\varphi})|-|z(\rho e^{i\varphi})|\leqslant|z(e^{i\varphi})-z(\rho e^{i\varphi})|.$$

This is less than $2(1-\rho)$ by the first mean-value theorem, since $|z'(w)|<2$ by (4.5). Thus for all θ,

$$1-|w(re^{i\theta})|=1-\rho>\tfrac{1}{2}(1-r).$$

On the other hand, we have

$$\begin{aligned}1-|w(re^{i\theta})|&=|w[\exp\{\psi(\theta)+i\theta\}]|-|w(re^{i\theta})|\\&\leqslant|w[\exp\{\psi(\theta)+i\theta\}]-w(re^{i\theta})|\\&<\frac{3}{2}\{\exp\psi(\theta)-r\},\end{aligned}$$

by the first mean-value theorem, since $|w'(z)|<3/2$ by (4.5). Hence by (4.4), for $|\theta-\theta_n|<1-r$,

*If z is on the boundary of D, then by $w'(z)$ we mean $\lim\{w(t)-w(z)\}/(t-z)$ as $t\to z$ with t in the closure of D.

$$1 - |w(re^{i\theta})| < \frac{3}{2}(2 - r - r) = 3(1 - r).$$

This proves Lemma 4.

5. In this section we make use of all the hypotheses of the theorem. Let D be as in §3, so that, in view of Lemma 2, the proof of the theorem will be complete when we have shown that a function $f(z)$ with the required properties exists.

Choose a sequence of integers $\{m_k\}$ such that, as $k \to \infty$,

$$\frac{\delta_{m_k}}{\delta_{m_k+1}} \to \infty, \tag{5.1}$$

$$\left(\sum_{m_k}^{\infty} \delta_j\right) \log \frac{1}{\delta_{m_k}} \to \infty. \tag{5.2}$$

Such a choice is possible by (2.4). Put

$$n_k = \left[\frac{1}{\delta_{m_k}}\right] \tag{5.3}$$

where the square brackets denote the integer part, and let

$$f(z) = \sum_{j=1}^{\infty} j^{-2}\{w(z)\}^{n_j}, \tag{5.4}$$

where $w(z)$ is as in §4. Since $w(z)$ is regular in D and $|w(z)| < 1$ there, it follows that $f(z)$ is regular and bounded in D.

Let

$$r_k = 1 - \delta_{m_k}, \tag{5.5}$$

and denote by E_k the union of the set of arcs

$$z = r_k e^{i\theta}, \quad |\theta - \theta_n| < \delta_{m_k} \qquad (n = 1, 2, 3, \ldots).$$

When $z \in E_k$ we have, by Lemma 4 and (5.5),

$$\tfrac{1}{2}\delta_{m_k} < 1 - |w(z)| < 3\delta_{m_k}.$$

Hence by (5.4), when $z \in E_k$,

$$\begin{aligned}\left|\frac{w(z)f'(z)}{w'(z)}\right| &= \left|\sum_{j=1}^{\infty} j^{-2} n_j \{w(z)\}^{n_j}\right| \\ &> k^{-2} n_k |w(z)|^{n_k} - \sum_{1}^{k-1} j^{-2} n_j |w(z)|^{n_j} - \sum_{k+1}^{\infty} j^{-2} n_j |w(z)|^{n_j} \\ &> k^{-2} n_k (1 - 3\delta_{m_k})^{n_k} - \sum_{1}^{k-1} j^{-2} n_j - \sum_{k+1}^{\infty} j^{-2} n_j (1 - \tfrac{1}{2}\delta_{m_k})^{n_k},\end{aligned}$$

that is,

$$\left|\frac{w(z)f'(z)}{w'(z)}\right| > k^{-2} n_k (1 - 3\delta_{m_k})^{n_k} - \Sigma_1 - \Sigma_2, \tag{5.6}$$

say.

By (5.3),

$$(1 - 3\,\delta_{m_k})^{n_k} \to e^{-3} \qquad (k \to \infty). \tag{5.7}$$

Moreover, by (5.1) and (5.3),

$$\sum\nolimits_1 = O(1)(k-1)^{-2} n_{k-1} = o(k^{-2} n_k). \tag{5.8}$$

Further, the simple inequality

$$(1 - x)^n < (nx)^{-2},$$

which holds for $0 < x < 1$ and $n > 0$, gives

$$(1 - \tfrac{1}{2}\,\delta_{m_k})^{n_j} < 4\, n_j^{-2} \delta_{m_k}^{-2} = O(n_j^{-2} n_k^2),$$

by (5.3). Hence by (5.1) and (5.3),

$$\begin{aligned} \sum\nolimits_2 &= O(n_k^2) \sum_{j=k+1}^{\infty} j^{-2} n_j^{-1} \\ &= O(k^{-2} n_k^2 n_{k+1}^{-1}), \end{aligned}$$

and so

$$\sum\nolimits_2 = o(k^{-2} n_k). \tag{5.9}$$

Combining (5.6), (5.7), (5.8), and (5.9), we get

$$\left| \frac{w(z) f'(z)}{w'(z)} \right| > 2\eta k^{-2} n_k,$$

for all large enough k, where $\eta > 0$ is a constant. But by (4.5) and the fact that $|w(z)| < 1$ we have $|w'(z)/w(z)| > \frac{1}{2}$, and so

$$|f'(z)| > \eta k^{-2} n_k \qquad (z \in E_k) \tag{5.10}$$

for all large enough k.

By the monotonicity of $\{\delta_n\}$ and the definition of E_k, the union of the sequence of *non-overlapping* arcs

$$z = r_k e^{i\theta},\ \theta_n < \theta < \theta_n + \delta_n,\ n = m_k, m_k + 1, m_k + 2, \ldots,$$

is contained in E_k. Hence the angular measure of E_k is at least

$$\sum_{m_k}^{\infty} \delta_j.$$

But

$$m(r_k, f') > \frac{1}{2\pi} \int \log^+ |\ldots| f'(r_k e^{i\theta})|\, d\theta,$$

the integral being taken over the set of θ in $(-\pi, \pi)$ such that $r_k e^{i\theta} \in E_k$. Hence by (5.10),

$$m(r_k, f') > \frac{1}{2\pi}\left(\sum_{j=m_k}^{\infty} \delta_j\right) \log (\eta k^{-2} n_k),$$

for all large enough k. It is easy to deduce from this, with (5.1), (5.2), and (5.3), that $m(r_k, f') \to \infty$ as $k \to \infty$. Since $m(r, f')$ is an increasing function of r (**4**, p. 8), it follows that $m(r, f') \to \infty$ as $r \to 1$. This proves the theorem.

6. It is easy to see from Theorem III of (**2**) that the condition (2.2) is essential for the truth of Hayman's theorem, and that some "order-of-contact condition," of the kind occurring in the definition of "properly contains," is also necessary. It is possible to construct examples giving more complete information about these hypotheses, but this does not seem worthwhile.

Professor Hayman has sent me the following remark. "As Professor A. A. Goldberg has kindly pointed out to me, Theorem III of (**1**) is in fact contained in a more general result of R. Nevanlinna (*Acta Soc. Sci. Fennicae, 50*, no. 12 (1925)), and this paper is also in other ways related to (**1**)."

References

1. W. K. Hayman, *On Nevanlinna's second theorem and extensions*, Rend. Circ. Mat. Palermo, Ser. II, *II* (1953).
2. P. B. Kennedy, *A property of bounded regular functions*, Proc. Roy. Irish Acad., *60*, sec. A, no. 2 (1959).
3. J. E. Littlewood, *Lectures on the Theory of Functions* (Oxford, 1944).
4. R. Nevanlinna, *Le Théorème de Picard-Borel et la Théorie des Fonctions Méromorphes* (Paris, 1929).
5. E. J. Specht, *Estimates on the mapping function and its derivatives in conformal mapping of nearly circular regions*, Trans. Amer. Math. Soc., *71* (1951), 183–96.

University College,
Cork, Ireland

SOME PRINCIPLES UNDERLYING THE CONSTRUCTION OF MEASURES

CHARLES A. HAYES, JR.

1. Introduction. Measures may be obtained from suitable non-negative valued functions in a number of ways. It is the purpose of this paper to present an abstract formulation of certain principles which may be used to construct measures, and to show that the various methods most frequently encountered in the literature are in fact all special applications of these principles.

The basic requirements A1–A4 are set forth in § 2, and it is there shown that a measure can be defined when these are fulfilled.* These requirements are satisfied in every case known to the writer. Usually, however, conditions stronger than A1–A4 hold, and it is these extra restrictions which yield information on the class of measurable sets and other matters. In §§ 3, 4, and 5 certain abstracts form of such restrictions are considered, and results are derived thereforom. The paper concludes with an analysis of how a number of measures occur as special cases of the theory.

We begin by stating a number of definitions and conventions which will be used throughout the paper.

If $\mathfrak{F}$ is any family of subsets of a set S, we agree to let $\cup \mathfrak{F}$ denote the *union* of $\mathfrak{F}$; that is, the set of all points which belong to at least one member of $\mathfrak{F}$. For a sequence of sets A we sometimes use the notation $\cup_{n=1}^{\infty} A_n$ to denote the union. If A and B are sets, then by $A - B$ we shall mean the set of those points which are in A but not in B. We agree to denote by $U(x)$ the set whose sole member is x. We further agree to let 0 denote the null set as well as zero.

When discussing metric spaces, the terms sphere, closed, open, Borel, etc., will be understood to have their customary meanings with reference to the metric of the space. If A and B are subsets of a metric space, the diameter of A will be denoted by diam A, with the convention that diam $0 = 0$, and the distance between A and B will be denoted by dist (A, B). We keep in mind that inf $0 = \infty$ and empty sums are zero.

We shall say that the *function* ϕ *measures* S if, and only if,

$$0 < \phi(A) < \sum_{\beta \in \mathfrak{F}} \phi(\beta)$$

Received September 10, 1958.

*These principles were formulated as an outgrowth of an observation by A. P. Morse, who noticed that a certain measure occurring in (2), discussed in 6.3 of the present paper, satisfies conditions which are special cases of the above.

whenever $\mathfrak{F}$ is a finite or countably infinite family and $A \subset \cup \mathfrak{F} \subset S$. This definition is equivalent to the usual properties, namely $\phi(0) = 0$; $\phi(A) \leq \phi(B)$ whenever $A \subset B \subset S$; and

$$\phi(\cup \mathfrak{F}) \leq \sum_{\beta \in \mathfrak{F}} \phi(\beta)$$

whenever $\mathfrak{F}$ is a finite or countably infinite family and $\cup \mathfrak{F} \subset S$.

The set $A \subset S$ is said to be *ϕ-measurable* if, and only if, ϕ measures S and

$$\phi(E) = \phi(A \cap E) + \phi((S - A) \cap E)$$

whenever $E \subset S$.

A family of sets $\mathfrak{F}$ is said to be a *ring* if, and only if, $(E \cup F) \in \mathfrak{F}$ and $(E - F) \in \mathfrak{F}$ whenever $E \in \mathfrak{F}$ and $F \in \mathfrak{F}$.

A family of sets $\mathfrak{F}$ is said to *cover* a set A if, and only if, $A \subset \cup \mathfrak{F}$.

We say that F is a *blanket* if, and only if, F is such a function that for x in its domain:

(i) S is a metric space;

(ii) $x \in S$ and $F(x)$ is a non-vacuous family of subsets of S;

(iii) $\operatorname{diam} \beta < \infty$ whenever $\beta \in F(x)$;

(iv) $\inf_{\beta \in F(x)} \operatorname{diam} (\beta \cup U(x)) = 0$.

If F is a blanket with domain A, then $\cup_{x \in A} F(x)$ is called the *spread* of F.

2. The general theory. At this point we introduce certain basic sets and functions which we shall understand to be fixed throughout the paper, along with the properties A1–A4 given below. In § 6, we will give certain examples which may serve to furnish motivation for the work to follow.

We let S and $\mathfrak{F}_0$ denote, respectively, a non-empty point set and a non-empty class of subsets* of S. It will turn out that S is our measure space. We let T denote another non-empty set which may or may not bear any relation to S.

We further fix a function $\mathbf{M}$ with the following properties:

A1. For each set $A \subset S$ and each point $t \in T$, $\mathbf{M}(A, t)$ is a collection (possibly empty) of finite or countably infinite subfamilies of $\mathfrak{F}_0$.

A2. $0 \in \mathbf{M}(0, t)$ for each $t \in T$.

A3. $\mathbf{M}(B, t) \subset \mathbf{M}(A, t)$ whenever $A \subset B \subset S$ and $t \in T$.

A4. If $t \in T$ and $\mathfrak{H}$ is any finite or countably infinite collection of subsets of S, then there exists a function H on $\mathfrak{H}$ to T such that if Q is any function on $\mathfrak{H}$ with $Q(\gamma) \in \mathbf{M}(\gamma, H(\gamma))$ whenever $\gamma \in \mathfrak{H}$, then

$$\cup_{\gamma \in \mathfrak{H}} Q(\gamma) \in \mathbf{M}(\cup \mathfrak{H}, t).$$

Finally, we fix a non-negative finite-valued function f whose domain is

*One could take $\mathfrak{F}_0$ as a class of subsets of a space other than S without affecting the validity of any results of § 2.

$\mathfrak{F}_0$. With f we associate the function $\bar{f}$ whose domain is the class of all subsets of S, so defined that whenever $A \subset S$,

$$\bar{f}(A) = \sup_{t \in T} \left\{ \mathfrak{G} \inf_{\mathfrak{G} \in \mathbf{M}(A,t)} \sum_{\gamma \in \mathfrak{G}} f(\gamma) \right\}.$$

We propose to investigate some of the properties of $\bar{f}$.

2.1. LEMMA. *If $A \subset S$, $\epsilon > 0$, and $t \in T$, then there exists a family $\mathfrak{R} \in \mathbf{M}(A,t)$ for which*

$$\sum_{\alpha \in \mathfrak{R}} f(\alpha) < \bar{f}(A) + \epsilon$$

Proof. Evidently

$$\mathfrak{G} \inf_{\mathfrak{G} \in \mathbf{M}(A,t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) < \bar{f}(A);$$

hence there exists a family $\mathfrak{R} \in \mathbf{M}(A,t)$ for which, as required,

$$\sum_{\alpha \in \mathfrak{R}} f(\alpha) < \left(\mathfrak{G} \inf_{\mathfrak{G} \in \mathbf{M}(A,t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) \right) + \epsilon < \bar{f}(A) + \epsilon.$$

2.2. THEOREM. *$\bar{f}$ measures S.*

Proof. From A2 above we clearly have

$$0 < \mathfrak{G} \inf_{\mathfrak{G} \in \mathbf{M}(0,t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) < \sum_{\alpha \in 0} f(\alpha) = 0$$

for each $t \in T$; consequently $\bar{f}(0) = 0$.

If $A \subset B \subset S$, then from A3, we see that for each $t \in T$,

$$0 < \mathfrak{G} \inf_{\mathfrak{G} \in \mathbf{M}(A,t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) < \mathfrak{G} \inf_{\mathfrak{G} \in \mathbf{M}(B,t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) < \bar{f}(B),$$

whence it follows that $0 < \bar{f}(A) < \bar{f}(B)$.

Finally, we consider a finite or countably infinite family $\mathfrak{H}$ of subsets of S. Given $\epsilon > 0$, we select any convenient positive-valued function η with domain $\mathfrak{H}$ for which

$$\sum_{\gamma \in \mathfrak{H}} \eta(\gamma) < \epsilon. \tag{1}$$

If $t \in T$, we may use A4 to select a function H on $\mathfrak{H}$ to T, enjoying the properties therein specified. Further, we may use Lemma 2.1 to find such a function Q that for each $\gamma \in \mathfrak{H}$,

$$Q(\gamma) \in \mathbf{M}(\gamma, H(\gamma)), \sum_{\alpha \in Q(\gamma)} f(\alpha) < \bar{f}(\gamma) + \eta(\gamma). \tag{2}$$

We let $\mathfrak{R} = \cup_{\gamma \in \mathfrak{H}} Q(\gamma)$. From (2) and the choice of H under A4 we infer that

$$\mathfrak{R} \in \mathbf{M}(\cup \mathfrak{H}, t). \tag{3}$$

Thus, using (1), (2), and (3), we see that

$$\mathfrak{G}\inf_{\mathfrak{G}\in\mathbf{M}(\cup\mathfrak{H},\,t)}\sum_{\alpha\in\mathfrak{G}}f(\alpha)<\sum_{\alpha\in\mathfrak{R}}f(\alpha)<\sum_{\gamma\in\mathfrak{H}}\left(\sum_{\alpha\in\mathfrak{Q}(\gamma)}f(\alpha)\right)$$
$$<\sum_{\gamma\in\mathfrak{H}}(\bar{f}(\gamma)+\eta(\gamma))<\sum_{\gamma\in\mathfrak{H}}\bar{f}(\gamma)+\epsilon.$$

From the arbitrary nature of both t and ϵ in this last relation, we conclude that

$$\bar{f}(\cup\mathfrak{H})<\sum_{\gamma\in\mathfrak{H}}f(\gamma),$$

thus completing the proof of the theorem.

3. A specialization of the general theory. In order to obtain further specific information concerning $\bar{f}$ we shall observe the effect on $\bar{f}$ when tee functions $\mathbf{M}$ and f are subjected to certain restrictions in addition to those already imposed.

For the remainder of this section we shall assume that T is directed by $\prec$ and that the following holds:

$A3'$. $\mathbf{M}(B,t_1)\subset\mathbf{M}(A,t_2)$ whenever $A\subset B\subset S$, $t_1\in T$, $t_2\in T$ and $t_2\prec t_1$.

Clearly, A3′ is a stronger restriction than $A3$. In this case it is easy to see that for any set $A\subset S$, $\bar{f}(A)$ may be expressed as a Moore-Smith limit. For any $t\in T$, we let

$$\bar{f}_t(A)=\mathfrak{G}\inf_{\mathfrak{G}\in\mathbf{M}(A,\,t)}\sum_{\alpha\in\mathfrak{G}}f(\alpha);$$

and then it follows that

$$\bar{f}(A)=\sup_{t\in T}\bar{f}_t(A)=\lim_{t\prec}f_t(A).$$

3.1. DEFINITION. *If $\mathfrak{G}$ is any family of sets and β is any set, then by $\mathfrak{G}\odot\beta$ we shall mean that family of sets $\mathfrak{G}'$ for which $\gamma\in\mathfrak{G}'$ if, and only if, $\gamma=\alpha\cap\beta$, where α is some member of $\mathfrak{G}$.*

3.2. DEFINITION. *If $A\subset S$, then we shall write $\bar{A}=S-A$.*

We are now ready to formulate a condition on $\mathbf{M}$ which bears on the nature of the class of $\bar{f}$-measurable sets.

A5. If $\beta\in\mathfrak{F}_0$, $A\subset S$, and $t\in T$, then $\mathfrak{G}\odot\beta\in\mathbf{M}(A\cap\beta,t)$ and $\mathfrak{G}\odot\bar{\beta}\in\mathbf{M}(A\cap\bar{\beta},t)$ whenever $\mathfrak{G}\in\mathbf{M}(A,t)$.

This involves possible restrictions on the class $\mathfrak{F}_0$.

3.3. DEFINITION. *We shall say that f is weakly quasi-additive on its domain $\mathfrak{F}_0$ if, and only if,*

(i) $\alpha \cap \beta \in \mathfrak{F}_0$ *and* $\alpha \cap \bar{\beta} \in \mathfrak{F}_0$ *whenever* $\alpha \in \mathfrak{F}_0$ *and* $\beta \in \mathfrak{F}_0$;
(ii) $f(\alpha) \geqslant f(\alpha \cap \beta) + f(\alpha \cap \bar{\beta})$ *whenever* $\alpha \in \mathfrak{F}_0$ *and* $\beta \in \mathfrak{F}_0$.

3.4. THEOREM. *If* $\mathbf{M}$ *satisfies A5 and* f *is weakly quasi-additive on* $\mathfrak{F}_0$, *then each member of* $\mathfrak{F}_0$ *is* $\bar{f}$*-measurable.*

Proof. We consider any set $A \subset S$, arbitrary members β and t of $\mathfrak{F}_0$ and T respectively, and an arbitrary member $\mathfrak{H}$ of $\mathbf{M}(A, t)$. Then, using A5 and Definition 2.3 we clearly have

$$\begin{aligned}(1)\qquad \sum_{\alpha \in \mathfrak{H}} f(\alpha) &\geqslant \sum_{\alpha \in \mathfrak{H}} f(\alpha \cap \beta) + \sum_{\alpha \in \mathfrak{H}} f(\alpha \cap \bar{\beta}) \\ &> \sum_{\gamma \in \mathfrak{H} \odot \beta} f(\gamma) + \sum_{\gamma' \in \mathfrak{H} \odot \bar{\beta}} f(\gamma') \\ &> \inf_{\mathfrak{G} \in \mathbf{M}(A \cap \beta,\, t)} \sum_{\gamma \in \mathfrak{G}} f(\gamma) + \inf_{\mathfrak{G}' \in \mathbf{M}(A \cap \bar{\beta},\, t)} \sum_{\gamma' \in \mathfrak{G}'} f(\gamma') \\ &= \bar{f}_t(A \cap \beta) + \bar{f}_t(A \cap \bar{\beta}).\end{aligned}$$

From the arbitrary nature of $\mathfrak{H}$ in (1) we obtain at once

$$(2)\qquad \bar{f}_t(A) \geqslant \bar{f}_t(A \cap \beta) + \bar{f}_t(A \cap \bar{\beta}).$$

Since t is arbitrary in (2), we obtain at once from the theory of Moore-Smith limits that

$$\begin{aligned}\bar{f}(A) &= \lim_{t,>} \bar{f}_t(A) \geqslant \lim_{t,>} \bar{f}_t(A \cap \beta) + \lim_{t,>} \bar{f}_t(A \cap \bar{\beta}) \\ &= f(A \cap \beta) + \bar{f}(A \cap \bar{\beta}),\end{aligned}$$

so that β is $\bar{f}$-measurable and the proof is complete.

4. Another form of specialization. We consider now another specialization of $\mathbf{M}$ and $\mathfrak{F}_0$ which is encountered in some applications. In such cases, f is not subjected to any requirements over and above those of § 2, except as the specialization just mentioned affects its domain $\mathfrak{F}_0$.

4.1. DEFINITION. The element $t \in T$ is said to be an ϵ-*approximator* for $\bar{f}(A)$ if, and only if, $A \subset S$ and

$$\bar{f}(A) < \sum_{\alpha \in \mathfrak{G}} f(\alpha) + \epsilon$$

whenever $\mathfrak{G} \in \mathbf{M}(A, t)$.

4.2. LEMMA. *If* $\epsilon > 0$, $A \subset S$, *and* $\bar{f}(A) < \infty$ *then there exists some* $t \in T$ *which is an* ϵ*-approximator for* $\bar{f}(A)$.

Proof. It follows from the definition of $\bar{f}$ in § 2 that there exists some $t \in T$ for which

$$\bar{f}(A) < \inf_{\mathfrak{G} \in \mathbf{M}(A,\, t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) + \epsilon,$$

from which it is clear that t is the desired element.

A6. There exists a relation R whose members are certain ordered pairs of subsets of S, with the following properties:

(i) whenever $A' \subset A \subset S$, $B' \subset B \subset S$, and $(A, B) \in R$, then $(A', B') \in R$;

(ii) if $A \subset S$, $B \subset S$, $C \subset S$, $(A, C) \in R$, and $(B, C) \in R$, then $(A \cup B, C) \in R$;

(iii) whenever $\bar{f}(A) < \infty$, $\bar{f}(B) < \infty$, $(A, B) \in R$, and $\epsilon > 0$, then there exist elements t, t', t'' of T such that t' and t'' are, respectively, ϵ-approximators for $\bar{f}(A)$ and $\bar{f}(B)$; in addition, whenever $\mathfrak{G} \in \mathbf{M}(A \cup B, t)$ then there exist two disjointed families $\mathfrak{G}'$ and $\mathfrak{G}''$ with $\mathfrak{G}' \in \mathbf{M}(A, t')$, $\mathfrak{G}'' \in \mathbf{M}(B, t'')$, such that $\mathfrak{G}' \cup \mathfrak{G}'' \subset \mathfrak{G}$.

4.3. THEOREM. *If* $\mathbf{M}$ *and* $\mathfrak{F}_0$ *satisfy* A6 *and* $(A, B) \in R$ *then* $\bar{f}(A \cup B) = \bar{f}(A) + \bar{f}(B)$.

Proof. Since $\bar{f}$ measures S by Theorem 2.2, we need prove only that $\bar{f}(A \cup B) \geq \bar{f}(A) + \bar{f}(B)$. Since this inequality clearly holds if either $\bar{f}(A)$ or $\bar{f}(B)$ is infinite, we may restrict ourselves to the case where $\bar{f}(A)$ and $\bar{f}(B)$ are both finite. We take $\epsilon > 0$, determine elements t, t', t'' of T in accordance with A6 (iii), and note that

$$\inf_{\mathfrak{H} \in \mathbf{M}(A \cup B,\, t)} \sum_{\alpha \in \mathfrak{H}} f(\alpha) < \bar{f}(A \cup B). \tag{1}$$

We select $\mathfrak{G} \in \mathbf{M}(A \cup B, t)$ so that

$$\inf_{\mathfrak{H} \in \mathbf{M}(A \cup B,\, t)} \sum_{\alpha \in \mathfrak{H}} f(\alpha) > \sum_{\alpha \in \mathfrak{G}} f(\alpha) - \epsilon. \tag{2}$$

Applying A6 (iii) we determine two disjointed families $\mathfrak{G}'$ and $\mathfrak{G}''$ with $\mathfrak{G}' \in \mathbf{M}(A, t')$, $\mathfrak{G}'' \in \mathbf{M}(B, t'')$, $\mathfrak{G}' \cup \mathfrak{G}'' \subset \mathfrak{G}$, put (1) and (2) together and use the nature of the choices of t' and t'' to obtain

$$\begin{aligned} \bar{f}(A \cup B) &> \sum_{\alpha \in \mathfrak{G}} f(\alpha) - \epsilon \\ &> \sum_{\alpha \in \mathfrak{G}'} f(\alpha) + \sum_{\alpha \in \mathfrak{G}'} f(\alpha) - \epsilon \\ &> (\bar{f}(A) - \epsilon) + (\bar{f}(B) - \epsilon) - \epsilon = \bar{f}(A) + \bar{f}(B) - 3\epsilon. \end{aligned} \tag{3}$$

The arbitrary nature of ϵ permits the desired conclusion.

A useful result which may be proved by simple induction based on A6 (ii) is stated in the following lemma.

4.4. LEMMA. *If* $B \subset S$, $A_i \subset S$, *and* $(A_i, B) \in R$ *for* $i = 1, 2, \ldots, n$, *then*

$$\left(\bigcup_{i=1}^{n} A_i, B \right) \in R.$$

We now state another useful result which follows by simple induction on Theorem 4.3 and Lemma 4.4.

4.5. THEOREM. *If* **M** *and* $\mathfrak{F}_0$ *satisfy A6,* $A_i \subset S$ *for* $i = 1, 2, \ldots, n$, *and* $(A_i, A_j) \in R$ *for each pair of positive integers* i *and* j *such that* $1 \leqslant i < j \leqslant n$, *then*

$$\bar{f}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \bar{f}(A_i).$$

4.6. LEMMA. *If* **M** *and* $\mathfrak{F}_0$ *satisfy A6,* $E \subset S$, $\beta \subset S$, *and there exists a positive integer* p *and a sequence of sets* Q *such that*

(a) $\beta = \bigcup_{i=1}^{\infty} Q_i$;

(b) $(Q_i, \bar{\beta}) \in R$ *for each positive integer* i;

(c) $(Q_i, Q_{i+pj}) \in R$ *for each positive integer* i *and* j,

then

$$\bar{f}(\beta \cap E) = \lim_n \bar{f}\left(\bigcup_{i=1}^{n} (Q_i \cap E)\right).$$

Proof. For each positive integer n, we let

$$A_n = \bigcup_{i=1}^{n} (Q_i \cap E) \subset \beta \cap E,$$

from which it follows that

$$\lim_n \bar{f}(A_n) \leqslant \bar{f}(\beta \cap E). \tag{1}$$

Now for each positive integer i, $0 \leqslant i \leqslant p - 1$, we consider the sets of the form Q_{i+pj}, $j = 1, 2, \ldots$. From (c) of the hypotheses, we see that $(Q_{i+pj}, Q_{i+pj'}) \in R$ whenever $1 \leqslant j < j'$, so that Theorem 4.5 yields

$$\sum_{j=1}^{n} \bar{f}(Q_{i+pj} \cap E) = \bar{f}\left(\bigcup_{j=1}^{n} (Q_{i+pj} \cap E)\right) \leqslant \bar{f}(A_{i+pn}) \leqslant \bar{f}(\beta \cap E). \tag{2}$$

In case

$$\sum_{j=1}^{\infty} \bar{f}(Q_{i+pj} \cap E) = \infty$$

for some i, $0 \leqslant i \subset p - 1$, then we see from (2) that $\lim_n \bar{f}(A_{i+pn}) = \infty$, hence $\lim_n \bar{f}(A_n) = \bar{f}(\beta \cap E) = \infty$, whence the desired result holds. If, however,

$$\sum_{j=1}^{\infty} \bar{f}(Q_{i+pj} \cap E) < \infty$$

for each i, $0 \leqslant i \leqslant p - 1$, then for anygiven $\epsilon > 0$ there exists such a positive integer N that for each i, $0 \leqslant i \leqslant p - 1$,

$$\sum_{j=N}^{\infty} \bar{f}(Q_{i+pj} \cap E) < \epsilon/p,$$

whence it follows that

$$\sum_{n=N_p}^{\infty} \bar{f}(Q_n \cap E) < \epsilon. \tag{3}$$

Since $\bar{f}$ measures S and

$$\beta \cap E = A_{N_p-1} \cup \bigcup_{i=N_p}^{\infty} (Q_i \cap E),$$

we have, using (1) and (3),

$$\begin{aligned} \bar{f}(\beta \cap E) &\leqslant \bar{f}(A_{N_p-1}) + \sum_{i=N_p}^{\infty} \bar{f}(Q_i \cap E) \\ &\leqslant \bar{f}(A_{N_p-1}) + \epsilon \leqslant \lim_n \bar{f}(A_n) + \epsilon \leqslant \bar{f}(\beta \cap E) + \epsilon. \end{aligned}$$

From this last relation and the arbitrary nature of ϵ, the desired result follows.

4.7. THEOREM. *Under the hypotheses of Lemma* 4.6, β *is* $\bar{f}$*-measurable.*

Proof. Let E denote an arbitrary subset of S. Using again the notation of Lemma 4.6, we showed there that

$$\lim_n \bar{f}(A_n) = \bar{f}(\beta \cap E), A_n \subset \beta \cap E \text{ for } n = 1, 2, \ldots. \tag{1}$$

From (b) of the hypotheses of Lemma 4.6, A6 (i), the definition of A_n, and Lemma 4.4 we have

$$(A_n, \bar{\beta} \cap E) \in R, \qquad n = 1, 2, \ldots. \tag{2}$$

Using (1), (2), and Theorem 4.3, we see that

$$\bar{f}(E) = \bar{f}(\beta \cap E) \cup (\bar{\beta} \cap E)) \geqslant \bar{f}(A_n \cup (\bar{\beta} \cap E)) = \bar{f}(A_n) + \bar{f}(\bar{\beta} \cap E)$$

for each positive integer n; applying (2) we have finally

$$\bar{f}(E) \geqslant \bar{f}(\beta \cap E) + \bar{f}(\bar{\beta} \cap E),$$

which establishes the $\bar{f}$-measurability of β.

5. Further considerations. Unrelated to questions of measurability but of interest in itself is the relationship between $\bar{f}(\beta)$ and $f(\beta)$ for $\beta \in \mathfrak{F}_0$. A relatively simple restriction on $\mathbf{M}$, valid in many applications and yielding information on this point is now given.

A7. $U(\beta) \in \mathbf{M}(\beta, t)$ whenever $\beta \in \mathfrak{F}_0$ and $t \in T$.

5.1. THEOREM. *If* $\mathbf{M}$ *satisfies* A7 *then* $\bar{f}(\beta) \leqslant f(\beta)$ *whenever* $\beta \in \mathfrak{F}_0$.

Proof. For each $t \in T$ we have

$$\mathfrak{G} \inf_{\mathfrak{M} \in \mathbf{M}(\beta, t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) \leqslant f(\beta);$$

thus $\bar{f}(\beta) \leqslant f(\beta)$.

Although **M** satisfies A1–A4 as well as A7 in the theorem just proved, only A1 and A7 and the definition of $\bar{f}$ enter the proof.

To reverse the inequality just proved requires relatively specific properties of both **M** and f, and there appears to be no clear-cut restriction valid in most applications which yields the desired result. We do obtain the following result, however.

5.2. THEOREM. If **M** *satisfies* A7 *then a necessary and sufficient condition that* $\bar{f}(\beta) = f(\beta)$ *for each* $\beta \in \mathfrak{F}_0$ *is that for each* $\epsilon > 0$, *there exists some* $t_0 \in T$ *for which*

$$f(\beta) \leqslant \sum_{\alpha \in \mathfrak{G}} f(\alpha) + \epsilon$$

whenever $\mathfrak{G} \in \mathbf{M}(\beta, t_0)$.

Proof. If the stated condition holds, we clearly have

$$f(\beta) \leqslant \inf_{\mathfrak{G} \in \mathbf{M}(\beta, t_0)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) + \epsilon \leqslant \bar{f}(\beta) + \epsilon; \tag{1}$$

since ϵ is arbitrary then $f(\beta) \leqslant \bar{f}(\beta)$. The reverse inequality follows from Theorem 5.1.

If $f(\beta) = \bar{f}(\beta)$ and $\epsilon > 0$, then by the definition of $\bar{f}(\beta)$, there exists $t_0 \in T$ for which, since f is finite-valued,

$$f(\beta) = \bar{f}(\beta) \leqslant \inf_{\mathfrak{G} \in \mathbf{M}(\beta, t_0)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) + \epsilon,$$

from which follows at once the stated condition.

6. Applications of the theory.

6.1. The most common application of the theory just set forth consists of the case where S, $\mathfrak{F}_0$, and f satisfy the conditions of § 2, and for any set $A \subset S$, $\bar{f}(A)$ is the infimum of all numbers of the form $\sum_{\alpha \in \mathfrak{G}} f(\alpha)$, where $\mathfrak{G}$ denotes a finite or countably infinite subfamily of $\mathfrak{F}_0$ which covers A. Thus defined, $\bar{f}$ may not immediately seem to be in conformity with the definition in § 2. However, we may take for T any set consisting of a single element, for example, $T = U(0)$, and so define **M** that for each set $A \subset S$ and each $t \in T$, $\mathbf{M}(A, t) = \mathbf{M}(A, 0)$ consists of all finite or countably infinite subfamilies of $\mathfrak{F}_0$ which cover A. It is easily verified that A1–A4 are satisfied; in A4 one takes $H(\gamma) = 0 \in T$ for each $\gamma \in \mathfrak{H}$. With this interpretation it is also easily seen that $\bar{f}$ as defined in § 2 agrees with the definition given more briefly above. It is clear that $U(\beta) \in \mathbf{M}(\beta, t)$ for each $t \in T$ and each $\beta \in \mathfrak{F}_0$, hence by Theorem 5.1, $\bar{f}(\beta) \leqslant f(\beta)$ whenever $\beta \in \mathfrak{F}_0$.

In case f is a measure* and $\mathfrak{F}_0$ is a ring of sets, then A3′ and A5 hold and

*Here we use the term *measure* as defined in (3, p. 30).

Theorem 3.4 assures us that $\mathfrak{F}_0$ is contained in the class of $\bar{f}$-measurable sets. Furthermore, the criterion formulated in Theorem 5.2 is fulfilled, and consequently $\bar{f}$ agrees with f on $\mathfrak{F}_0$.

6.2. We now consider a second well-known method of defining a measure, where S is a metric space, $\mathfrak{F}_0$ is an arbitrary non-empty family of subsets of S, and f satisfies the requirements of § 2. Here we take T to be the set of all positive real numbers. For each set $A \subset S$ and each $t \in T$ we define $\mathbf{M}(A, t)$ to be the collection (possibly empty) of all finite or countably infinite subfamilies of $\mathfrak{F}_0$ which cover A, each of whose members is of diameter less than t. It is easily checked that A1–A4 are satisfied; in A4 one may take $H(\gamma) = t$ for each $\gamma \in \mathfrak{H}$. One may so define $\bar{f}_t$ for a given $t \in T$ that for each $A \subset S$

$$\bar{f}_t(A) = \inf_{\mathfrak{G} \in \mathbf{M}(A, t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha),$$

whence in accordance with the definition in § 2,

$$\bar{f}(A) = \sup_{t \in T} \bar{f}_t(A).$$

It is easily checked that for any given $t \in T$, $\bar{f}_t$ measures S; in fact, with t fixed, $\bar{f}_t$ is itself merely an example of the kind discussed in 6.1, with the added restriction on the covering families that their members be of diameter less than t. Also, in this case $\mathbf{M}$ satisfies condition A3′ of § 3 with $t_2 \prec t_1$ if, and only if, $t_2 \geqslant t_1$, and accordingly for any set $A \subset S$ we may write

$$\bar{f}(A) = \sup_{t \in T} \bar{f}_t(A) = \lim_{t \to 0+} \bar{f}_t(A),$$

which is the usual form of representation for $\bar{f}$.

If we introduce the relation

$$R = \{(A, B) : A \subset S, B \subset S, \text{ and } \operatorname{dist}(A, B) > 0\},$$

then A6 is valid for $\mathbf{M}$ and $\mathfrak{F}_0$. For A6 (i) and (ii) clearly hold. If $\bar{f}(A) < \infty$, $\bar{f}(B) < \infty$, $\epsilon > 0$, and $(A, B) \in R$, then one can evidently find $t \in T$ such that

$$0 < t < \operatorname{dist}(A, B),\ \bar{f}_t(A) > \bar{f}(A) - \epsilon,\ \bar{f}_t(B) > \bar{f}(B) - \epsilon,$$

from which t is seen to be an ϵ-approximator for both $\bar{f}(A)$ and $\bar{f}(B)$. If $\mathfrak{G} \in \mathbf{M}(A \cup B, t)$, we may define $\mathfrak{G}'$ as the subfamily of $\mathfrak{G}$ whose members intersect A, $\mathfrak{G}''$ as the remainder of $\mathfrak{G}$, and then $\mathfrak{G}' \in \mathbf{M}(A, t)$, $\mathfrak{G}'' \in \mathbf{M}(B, t)$, so that A6 (iii) holds.

If β is any open set and i is any positive integer, we may define $Q_i = \{x : 1/(1+i) \leqslant \operatorname{dist}(x, \beta) < 1/i\}$, from which it follows easily that the hypotheses (a), (b), and (c) of Lemma 4.6 hold, with $p = 2$. Thus Theorem 4.7 applies and all open sets are $\bar{f}$-measurable, thus so are all Borel sets as well.

As a special example of the above, we mention the case where $\mathfrak{F}_0$ consists of all open sets in S, q is a fixed positive number, and

$$f(\beta) = (\text{diam } \beta)^q$$

whenever $\beta \in \mathfrak{F}_0$. The resulting $\bar{f}$ in this case is Hausdorff q-dimensional measure; all Borel subsets of S are $\bar{f}$-measurable. Since A7 holds only for the null set and sets consisting of a single point, one cannot expect to make any general statement concerning the relative magnitudes of $f(\beta)$ and $\bar{f}(\beta)$ when $\beta \in \mathfrak{F}_0$.

Lebesgue measure in n-dimensional Euclidean space may be regarded as a further example of the above. One may take $\mathfrak{F}_0$ as the class of all open intervals in S, and f as that function on $\mathfrak{F}_0$ whose values are the hypervolumes of members of $\mathfrak{F}_0$. As above, one may first define $\bar{f}_t$ for each $t > 0$ and then take

$$\bar{f} = \lim_{t \to 0+} \bar{f}_t.$$

Due to the nature of Euclidean space, $\bar{f}_t$ does not depend upon the value of t; thus $\bar{f} = \bar{f}_\infty$. However $\bar{f}_\infty$ is clearly a measure-function of the type discussed in 6.1, with $\bar{f}_\infty(\beta) \leqslant f(\beta)$ for each $\beta \in \mathfrak{F}_0$. Furthermore $\bar{f}_\infty$ is in conformity with the usual definition of Lebesgue outer measure. Thus $\bar{f}$ is ordinary Lebesgue outer measure and $\bar{f}(\beta) \leqslant f(\beta)$ for each $\beta \in \mathfrak{F}_0$. Since f satisfies the criterion of Theorem 5.2, we have in fact $\bar{f}(\beta) = f(\beta)$ for each $\beta \in \mathfrak{F}_0$. From above we see that all members of $\mathfrak{F}_0$ are $\bar{f}$-measurable, and consequently so are all Borel subsets of S.

6.3. A somewhat more complicated situation occurs in the generation of Borel measures. Here* we have a locally compact Hausdorff space S, the class $\mathfrak{C}$ of all compact subsets of S, and the class $\mathfrak{U}$ of all open sets belonging to $\mathfrak{C}$. The function λ is a *content* defined on $\mathfrak{C}$, and from λ is derived the function λ_* on $\mathfrak{U}$, so defined that for $G \in \mathfrak{U}$,

$$\lambda_*(G) = \sup_{G \supset C, \mathfrak{G}} \lambda(C).$$

It turns out that λ^* is subadditive on $\mathfrak{U}$ From λ^* is derived the outer measure μ^*, where for $A \subset S$,

$$\mu^*(A) = \inf_{A \subset G, \mathfrak{U}} \lambda_*(G).$$

To show that μ^* may be interpreted as an application of our general theory, we take $\mathfrak{F}_0 = \mathfrak{U}$, $T = U(0)$; for each set $A \subset S$ and $t = 0 \in T$ we take $\mathbf{M}(A, t)$ as the class of all finite or countably infinite subfamilies $\mathfrak{G}$ of $\mathfrak{U}$ which cover A. As in 6.1 it follows that A1–A4 are satisfied. From the subadditivity of λ_* in $\mathfrak{U}$ and the fact that $U(G) \in \mathbf{M}(A, 0)$ for each $G \in \mathfrak{U}$ such that $A \subset G$, it follows readily that

*The notation and terminology of (3, pp. 231–6) are used in this section.

$$\mu^*(A) = \inf_{A \subset G, G \in \mathfrak{U}} \lambda_*(G) = \mathfrak{G} \inf_{\in \mathbf{M}(A,t)} \sum_{\alpha \in \mathfrak{G}} \lambda_*(\alpha)$$

$$= \sup_{t \in T} \mathfrak{G} \inf_{\in \mathbf{M}(A,t)} \sum_{\alpha \in \mathfrak{G}} \lambda_*(\alpha),$$

so that μ^* is an outer measure by virtue of Theorem 2.2. Since A5 and the hypotheses of Theorem 5.2 are valid we have $\mu^*(G) = \lambda_*(G)$ whenever $G \in \mathfrak{U}$. In this example, the measurability of the Borel sets and the fact that μ^* induces a regular measure μ follow from the particular properties of λ and are not taken up here.

6.4. Other applications arise in the theory of differentiation of set functions. In these situations, $\mathfrak{F}_0$ is usually the spread of a blanket F, and f satisfies conditions considerably stronger than those of § 2. We let ϕ denote a function which measures S. For each $A \subset S$ one may define $\bar{f}(A)$ as the infimum of numbers of the form $\sum_{\alpha \in \mathfrak{G}} f(\alpha)$, where $\mathfrak{G}$ denotes a finite or countably infinite subset of $\mathfrak{F}_0$ for which $\phi(A - \cup \mathfrak{G}) = 0$ (that is, $\mathfrak{G}$ is a 0-covering of A). By making a rather obvious modification in the definition of M as given in 6.1, it is easy to check that the conditions A1–A4 are satisfied and $\bar{f}$ measures S. Another function $\bar{f}$ arises in some situations. Here for each fixed $t > 0$ and each set $A \subset S$ we define $\bar{f}_t(A)$ as the infimum of numbers of the form $\sum_{\alpha \in \mathfrak{G}} f(\alpha)$, where $\mathfrak{G}$ is a finite or countably infinite subset of $\mathfrak{F}_0$ for which $\phi(A - \cup \mathfrak{G}) < t$ (that is, $\mathfrak{G}$ is a t-covering of A). Then one defines

$$\bar{f}(A) = \lim_{t \to 0+} \bar{f}_t(A).$$

If one takes for T the set of all positive real numbers, and so defines $\mathbf{M}$ that for each set $A \subset S$ and each $t \in T$, $\mathbf{M}(A, t)$ is the collection of all t-coverings of A, then it is not difficult to show that $\mathbf{M}$ satisfies A1–A4. In verifying A4, when $t \in T$ and $\mathfrak{H}$ are given, one chooses a positive-valued function H such that $\sum_{\gamma \in \mathfrak{H}} H(\gamma) < t$. If $\gamma \in \mathfrak{H}$ we choose any family $Q(\gamma) \in \mathbf{M}(\gamma, H(\gamma))$ and let $\mathfrak{Q} = \cup_{\gamma \in \mathfrak{H}} Q(\gamma)$, then $\phi(\gamma - \cup Q(\gamma)) < H(\gamma)$ whenever $\gamma \in \mathfrak{H}$ and

$$\phi(\cup \mathfrak{H} - \cup \mathfrak{Q}) \leqslant \phi\left(\cup_{\gamma \in \mathfrak{H}} (\gamma - \cup Q(\gamma))\right) \leqslant \sum_{\gamma \in \mathfrak{H}} \phi(\gamma - Q(\gamma)) < \sum_{\gamma \in \mathfrak{H}} H(\gamma) < t;$$

thus $\mathfrak{Q} \in \mathbf{M}(\cup \mathfrak{H}, t)$ and A4 holds.

In this example, condition A3′ of § 3 is again valid, which makes possible the equation

$$\bar{f}(A) = \lim_{t \to 0+} f_t(A) = \sup_{t \in T} \left\{ \mathfrak{G} \inf_{\in \mathbf{M}(A,t)} \sum_{\alpha \in \mathfrak{G}} f(\alpha) \right\}.$$

In the above, applications to differentiation theory A7 holds, so that by Theorem 5.1, $\bar{f}(\beta) \leqslant f(\beta)$ whenever $\beta \in \mathfrak{F}_0$.

As was mentioned earlier, f and F are usually subject to various technical requirements which make it possible to use the more specialized parts of the above theory to obtain information concerning the class of $\bar{f}$-measurable sets and other matters. These are too lengthy for consideration here.

References

1. A. P. Morse, *A theory of covering and differentiation*, Trans. Amer. Math. Soc., **55** (1944), 205–35.
2. C. A. Hayes, *Differentiation with respect to ϕ-pseudo-strong blankets and related problems*, Proc. Amer. Math. Soc., **3**, (1952), 283–96.
3. P. R. Halmos, *Measure Theory* (New York, 1950).

University of California,
Davis, California

PENCILS OF NULL POLARITIES

SEYMOUR SCHUSTER

1. Introduction. A theorem due to von Staudt states that a null polarity in complex projective space of three dimensions is determined by a self-polar skew pentagon. By allowing an element of the self-polar pentagon to vary in a suitable manner we can arrive at a family of ∞^1 null polarities, which we term a *pencil of null polarities*. Each polarity of the pencil distinguishes a linear complex as the class of self-polar lines. Thus, associated with the pencil is a family of ∞^1 linear complexes, which we term a *pencil of linear complexes*.

It is the purpose of this paper to continue an earlier investigation of pencils of polarities **(2)**, by applying analogous techniques to the study of pencils of null polarities and pencils of linear complexes.

Since it develops that the lines common to all linear complexes of a pencil are the lines of a linear congruence, the central question has been: How many of the different types of linear congruences can be achieved in this manner? Happily, it can be reported that the classification of pencils of null polarities yields all of the three types of linear congruences **(4**, pp. 140-141**)**.

We conclude with some remarks concerning such pencils in real projective space.

2. Basic notions and constructions. Our basic configuration (see Figure 1) shall be the skew pentagon $PQRST$, where we designate certain planes by Greek letters as follows:

$$TPQ = \pi,\ PQR = \eta,\ QRS = \rho,\ RST = \sigma,\ STP = \tau.$$

The pentagon will be called *complete* if no four vertices are coplanar, and *self-polar* with respect to a null polarity Π if Π makes the following correspondence:

$$P \to \pi,\ Q \to \eta,\ R \to \rho,\ S \to \sigma,\ T \to \tau.$$

The correlation Γ which distinguishes a single line l in such a manner that Γ maps each point X (not on l) into plane Xl, and each plane χ (not through l) into point $\chi \cdot l$, is clearly singular. We call Γ a *special null polarity* with *directrix* l. (In this connotation some authors might object to use of the word "polarity" since Γ is not $1-1$.) The points of l and the planes through l do not belong to the domain of definition of Γ.

VON STAUDT'S THEOREM. *If $PQRST$ is a complete pentagon, the correlation*

$$P \to \pi,\ Q \to \eta,\ R \to \rho,\ S \to \sigma,\ T \to \tau$$

Received January 6, 1958.

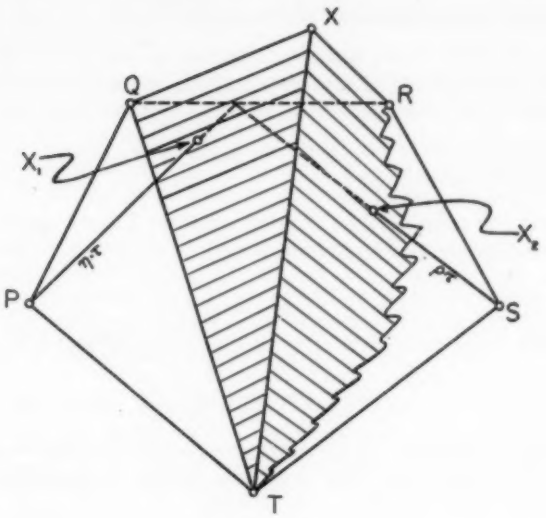


Figure 1

is a null polarity; that is, a complete self-polar pentagon determines a null polarity.

Since much of our work shall deal with the above correlation when the pentagon is not complete, we state the following: *If four vertices, say P, Q, R, and T are coplanar (with no three collinear), then* $\pi = \eta$, *and the correspondence*

$$P \to \pi,\ Q \to \eta,\ \rho \to R,\ S \to \sigma,\ \tau \to T$$

is a special null polarity with directrix RT. The verification of this fact is quite simple.

THEOREM 1. *Let X be a point in general position. The polar of X in the null polarity defined by the self-polar pentagon PQRST is*

$$\chi = X(XQT\cdot\eta\cdot\tau)(XRT\cdot\rho\cdot\tau). \tag{1}$$

Proof. (See Figure 1.) The pole of every plane through QT is on $\eta\cdot\tau$. Therefore, the pole of XQT is the point $X_1 = XQT\cdot\eta\cdot\tau$. Similarly, the pole of XRT is the point $X_2 = XRT\cdot\rho\cdot\tau$. Thus, the three points X, X_1 and X_2 lie on χ, and determine it according to (1).

In the event that the null polarity is special, (1) reduces to the simple expression of χ as the join of X with the directrix.

The dual procedure yields the construction for the pole of an arbitrary plane.

3. Pencils of null polarities.

Definition. A pencil of null polarities is the set of null polarities defined by the self-polar pentagon $PQRST$, where P, Q, R, and S are all fixed, while T varies on a line t (not through any other vertices) in its fixed polar plane τ.

It follows from the definition that plane π and/or σ also vary in axal pencils about PQ and RS, respectively.

THEOREM 2. *The polar planes χ of any fixed point X, with respect to a pencil of null polarities, form a pencil of planes.*

Proof. Referring to Figure 1, we call $X_1 = XQT\cdot\eta\cdot\tau$ and $X_2 = XRT\cdot\rho\cdot\tau$. Then, by Theorem 1, $\chi = X\,X_1X_2$. Let T vary on t, and consider the projectivity which maps the points of $\eta\cdot\tau$ onto the points of $\rho\cdot\tau$ as follows:

$$X_1 \overset{P}{\overline{\wedge}} QR\cdot\eta\cdot\tau \overset{S}{\overline{\wedge}} X_2$$

This projectivity possesses an invariant point, which arises when T is on plane XQR. Thus, the projectivity is a perspectivity. If we call the centre of perspectivity 0, then χ is always on the line XO.

We classify null polarities into three types, according to whether the line t meets none, one or two of the fixed edges of the pentagon. Precisely,

(i) P, Q, R, and S non-coplanar, and t in general position (not meeting any of the fixed edges of the defining pentagon).

(ii) P, Q, R, and S non-coplanar, but t meets exactly one fixed edge of the pentagon. It follows from the definition of a pencil (τ being fixed) that the single edge which t meets must be QR.

(iii) The line t meets two of the fixed edges of the defining pentagon.

It is easy to see that any other degeneracy of the pentagon, or any other position of t yields a pencil equivalent to one of the above types. We shall refer to these as the *general*, *parabolic*, and *degenerate systems*, respectively. (The justification for these names will be seen in §4.)

The general system. T varies on the line t not through any of the fixed lines of the pentagon. Thus, there are two distinct positions of T, in the plane η and in the plane ρ, which yield special null polarities. Although it may appear that further special null polarities arise when T is coplanar with P, R, and S, or similarly when T is coplanar with P, Q, and S, this is not the case. For, in these two instances it is possible to choose an alternate position anywhere on $\pi\cdot\sigma$ for the fifth point T of the self-polar pentagon; thus, showing that the polarity is, in fact, not special. Therefore, *there are exactly two special null polarities in a general system.*

THEOREM 3. *If two null polarities belong to the same general system, their product is a general axal homography* (1, 385); *conversely, every general axal*

homography can be expressed as the product to two null polarities belonging to the same general system.

Proof. The product of two such polarities leaves fixed the following elements: P, Q, R, S, η, ρ, τ, and t. Hence, $\tau \cdot QR$ is also fixed giving three invariant points on the line QR. It follows that QR is pointwise invariant, providing a point-axis for the homography. The tangential-axis is t, with the fixed points $t \cdot \eta$ and $t \cdot \rho$. Thus, the conditions for the general axal homography are established.

For the converse let the homography have QR as its point-axis and MN as its tangential-axis, with the collineation on the tangential axis defined by the projectivity $MNT_1 \overline{\wedge} MNT_2$. Then the given homography is the product of the two null polarities $PQRST_1$ and $PQRST_2$, where P and Q are arbitrary points on QRM and QRN, respectively.

The parabolic system. This pencil is defined by having the line t, the locus of T, meet the opposite edge QR. The position of $T = t \cdot QR$ yields a special null polarity, while all other positions of T yield non-special null polarities. Thus, *a parabolic system possesses exactly one special null polarity.*

THEOREM 4. *If two null polarities belong to the same parabolic system, their product is a biaxal homography* (**1**, pp. 385-386). *Conversely, every biaxal homography can be expressed as the product of two null polarities belonging to the same parabolic system.*

Proof. As in the proof of Theorem 3, the line QR is pointwise invariant under the product, since Q, R, and $t \cdot QR$ are three invariant points on the line. Further, the line PS is also pointwise invariant under the product, since P, S, and $t \cdot PS$ are three invariant points on that line. A biaxal homography is the projective transformation characterized by two such lines.

For a proof of the converse, let the homography be determined by the axes PS and QR, and a pair of corresponding points A and B (see Figure 2). Let T_1 be a point in general position, and call

$$\tau = PST_1 \quad \text{and} \quad t = \tau \cdot QRT_1$$

The null polarity $\Pi_1 = PQRST_1$ belongs to a parabolic system with T_1 varying on t. We shall determine another polarity of this system to satisfy the requirements of the theorem. We call

$$A_1 = AQT_1 \cdot \eta \cdot \tau, \quad A_2 = ART_1 \cdot \rho \cdot \tau \quad \text{and} \quad \alpha = A\,A_1A_2.$$

Then Π_1 maps $A \rightarrow \alpha$. Since τ is fixed for all members of the parabolic system, we may consider BA_1A_2 as the polar plane of B under one of the polarities in the system. To determine this polarity explicitly, we follow the converse construction calling $T_2 = BQA_1 \cdot t (= BRA_2 \cdot t)$. The desired polarity is then $\Pi_2 = PQRST_2$. The product $\Pi_1\Pi_2$ yields the original biaxal homography.

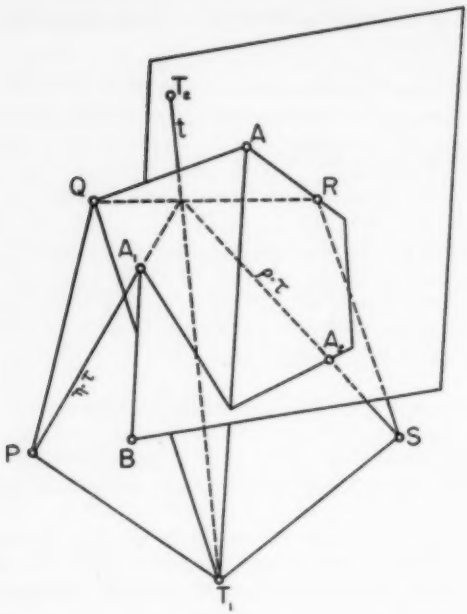


Figure 2

The degenerate system. To fix ideas let t meet edges PQ and QR. Thus, the points P, Q, R and T are always coplanar forcing every null polarity of the pencil to be special with directrix RT. The fixed plane τ is not defined in the degenerate system, sparing us the need of having S coplanar with the other four points, the consequence of which would be complete degeneracy with no null polarities defined. We may therefore conclude that *all the polarities of a degenerate system are special with their directrices forming a flat pencil of lines.*

4. Pencils of linear complexes. The self-polar lines of each null polarity of a pencil form a linear complex. The ∞^1 linear complexes which arise from a pencil are referred to as a *pencil of linear complexes* (**2**, pp. 92-93; **5**, pp. 332-333). Theorem 2 implies that each point X may be associated with a line x, called *the axis of* X, which is self-polar for all polarities in the pencil. (The axis of X is precisely the line OX mentioned in the proof of Theorem 2.) Further, every point of x is associated with the same axis. Thus, there are ∞^2 axes, and we may state

THEOREM 5. *The set of lines self-polar for all polarities of a pencil form a linear congruence.*

There are exactly three types of linear congruences in complex projective space (3, pp. 140-141). Our main result is that all three can be achieved as the axes of the three types of pencils of null polarities. More precisely, we state

THEOREM 6. (i) *The axes of a general pencil form a general linear congruence.*
(ii) *The axes of a parabolic pencil form a parabolic linear congruence.*
(iii) *The axes of a degenerate pencil form a degenerate linear congruence.*

Proof. (i) Each linear complex of the pencil has among its self-polar lines a flat pencil in η with vertex Q and a flat pencil in ρ with vertex R. Thus, the lines $R(t\cdot\eta)$ and $Q(t\cdot\rho)$ are the directrices of the linear congruence of axes. Further, $t\cdot\eta \neq t\cdot\rho$, from which it follows that the directrices are skew. Hence, the axes form a general linear congruence.

It is interesting to note that the line t is a member of the congruence, so that we could deduce—independently of the definition—that the plane τ must be fixed in a general pencil.

(ii) The two directrices of (i) become coincident (with QR), and all the axes meet QR. Since the polar plane of $QR\cdot t$ is QRT in every polarity of the pencil, we know that t is also an axis. Hence, we have the parabolic linear congruence established by the four linearly independent axes PQ, QR, RS, and t.

(iii) Let the degenerate system be defined by having t meet both PQ and QR. In §2, we saw that all the linear complexes of this pencil are special with directrix RT. Hence, all the lines of η belong to the linear congruence of lines which are common to all the linear complexes of the pencil.

Now consider a point X in general position. Its polar plane is XRT. Therefore, XR is self-polar in every polarity under consideration. The desired degenerate congruence is then established as consisting of the bundle of lines with vertex R plus the set of all lines in η.

Before concluding we turn attention to pencils of linear complexes in *real* projective space. In this case the general linear congruence is classified as hyperbolic or elliptic according as its directrices are proper (real) or improper (2, p. 93; 5, pp. 315-318). We therefore consider whether both the hyperbolic and the elliptic congruences can be attained as the axes of different types of general pencils.

An examination of the proof of Theorem 6 (i) shows that nothing is altered by the condition that $PQRST$ be a real pentagon. The congruence of axes has as its directrices the lines $R(t\cdot\eta)$ and $Q(t\cdot\rho)$, both of which are real. Hence, *a hyperbolic congruence results.* However, the elliptic congruence is unattainable for the very reason that lines $R(t\cdot\eta)$ and $Q(t\cdot\rho)$ are real lines (and are always the directrices). A further argument which also easily establishes the unattainability of the elliptic congruence (when P, Q, R, S, and T are real) is the following: An elliptic congruence is generated by four linearly

independent skew lines, such that no one of them meets the regulus containing the other three in a proper point (**5**, p. 315). Yet our general system —indeed, all of our cases—always include three non-skew linearly independent lines PQ, QR, and RS, as generators. Hence *it is impossible that the congruence of axes be elliptic.*

References

1. R. Baldus, *Zur Klassification der ebenen und räumlichen Kollineationen*, Sitz. Bayerischen Akad. (1928), 375–395.
2. H. S. M. Coxeter, *Non-euclidean geometry* (3rd ed.; Toronto 1957).
3. S. Schuster, *Pencils of polarities in projective space*, Can. J. Math., *8* (1956), 119–144.
4. J. A. Todd, *Projective and analytical geometry* (London, 1947).
5. O. Veblen and J. W. Young, *Projective geometry*, vol. 1 (Boston, 1910).

Carleton College
Northfield, Minnesota

ON THE SYMMETRY OF CUBIC GRAPHS

W. T. TUTTE

1. Introduction. Let G be a connected finite graph in which each edge has two distinct ends and no two distinct edges have the same pair of ends. We suppose further that G is *cubic*, that is, each vertex is incident with just three edges.

An *s-path* in G, where s is any positive integer, is a sequence $S = (v_0, v_1, \ldots, v_s)$ of $s + 1$ vertices of G, not necessarily all distinct, which satisfies the following two conditions:

(i) Any three consecutive terms of S are distinct.

(ii) Any two consecutive terms of S are the two ends of some edge of G.

If these conditions hold we call v_0 the *head* and v_s the *tail* of S.

An *automorphism* of G is a 1-1 mapping f of the set $V(G)$ of vertices of G onto itself such that fv and fw are the two ends of an edge of G if and only if v and w are the two ends of an edge of G. The automorphisms of G constitute a group $\mathbf{A}(G)$.

If $S = (v_0, v_1, \ldots, v_s)$ is any s-path of G we write fS for the s-path $(fv_0, fv_1, \ldots, fv_s)$, for each $f \in \mathbf{A}(G)$. We say G is *s-regular* if for each ordered pair $\{S, T\}$ of s-paths of G, not necessarily distinct, there is a unique element f of $\mathbf{A}(G)$ such that $T = fS$. Our main object in this paper is the proof of the following

THEOREM. *Suppose all the oriented edges of G are equivalent under $\mathbf{A}(G)$. Then there exists a positive integer s such that G is s-regular.*

As is explained in **(3)** the s-regular cubic graphs can be divided into two classes according to the nature of the group $\mathbf{A}(G)$. The first of these classes is discussed in the main paper of **(3)** and the second in the Addendum. It is shown that s is at most 5 for the first class and at most 4 for the second. Examples of graphs, of the first class only, are given for the values 2, 3, 4, and 5 of s. Other examples of the first class are given in **(1)**. In **(2)**, Frucht describes a 1-regular graph, without determining to which of the two classes it belongs. In §3 of the present paper we show that all 1-regular cubic graphs belong to the second class. Accordingly Frucht's graph is the first known member of this class.

2. Proof of the theorem. Let $S = (v_0, v_1, \ldots, v_s)$ be any s-path of G. The edges incident with v_0 join it to just two vertices, w and w' say, other than v_1.

Received January 22, 1958.

Similarly the edges incident with v_s join it to just two vertices, x and x' say, other than v_{s-1}. We call the s-paths $(w, v_0, v_1, \ldots, v_{s-1})$ and $(w', v_0, v_1, \ldots, v_{s-1})$ the *successors* of S. Similarly we call $(v_1, v_2, \ldots, v_s, x)$ and $(v_1, v_2, \ldots, v_s, x')$ the *predecessors* of S.

An s-path S' of G is *accessible* from S if there is a finite sequence $(S_1, S_2, \ldots, S_k)$ of s-paths of G satisfying the following conditions.

(i) $S_1 = S$ and $S_k = S'$.

(ii) For $1 < i < k$, S_{i+1} is either a predecessor or a successor of S_i in G.

(2.1) *Any s-path of G is accessible from any other (with the same value of s).*

Proof. Let S be any s-path of G. Let $\mathbf{W}$ be the class of all s-paths of G accessible from S. Then $S \in \mathbf{W}$.

Let V be the class of all vertices of G belonging to at least one member of $\mathbf{W}$. If $V(G) - V$ is not null then, by the connection of G we can find an edge E of G having one end p in $V(G) - V$ and one end q in V. Now q belongs to some $U_1 \in \mathbf{W}$. Starting with U_1 and taking predecessors as often as necessary we obtain $U_2 \in \mathbf{W}$ having q as its head. But then p is the head of a successor of U_2, contrary to the definition of p. We deduce that $V = V(G)$.

Let $S' = (w_0, w_1, \ldots, w_s)$ be any s-path of G. By the result just proved there exists $Z_1 \in \mathbf{W}$ such that w_s is a vertex of Z_1. From Z_1 by taking successors we obtain $Z_2 \in \mathbf{W}$ having w_s as its tail. From Z_2 by taking predecessors we can obtain $Z_3 \in \mathbf{W}$ with the following properties: The head of Z_3 is w_s and the second term of Z_3 is not w_{s-1}. Accordingly S' can be obtained from Z_3 by taking successors. Hence $S' \in \mathbf{W}$.

(2.2) *Suppose there is a positive integer s such that all the s-paths of G are equivalent under $\mathbf{A}(G)$ but G is not s-regular. Then for each s-path S of G there exists $f \in \mathbf{A}(G)$ such that $fS = S$ and f interchanges the two successors of S.*

Proof. Let S be any s-path of G, with successors T and T'. Since G is not s-regular there are s-paths Z_1 and Z_2 and distinct elements x and y of $\mathbf{A}(G)$ such that $Z_1 = xZ_2$ and $Z_1 = yZ_2$. There exists $z \in \mathbf{A}(G)$ such that $Z_1 = zS$. Write $g = z^{-1}xy^{-1}z$. Then $gS = S$ but g is not the identical automorphism of G.

We now show that there is an s-path S_1 of G such that S_1, but no successor of S_1, is invariant under g. For suppose not. Then if Z is any s-path invariant under g one of the successors of Z is also invariant under g. Hence both successors of Z are invariant under g. Let Z^{-1} be the s-path obtained from Z by reversing the order of the vertices. Then Z^{-1} and its successors are invariant under g. Hence the predecessors of Z are invariant under g, since they are obtained from the successors of Z^{-1} by reversing the order of vertices. Hence each s-path of G is invariant under g, by (2.1). This is impossible since g is not the identical automorphism of G.

Let the successors of S_1 be T_1 and T_1'. Let h be one of the elements of $\mathbf{A}(G)$ satisfying $hS = S_1$. Then h maps the successors of S onto the successors of S_1. We may adjust the notation so that $hT = T_1$ and $hT' = T_1'$. Since no successor of S_1 is invariant under g we have $gT_1 = T_1'$ and $gT_1' = T_1$. Write $f = h^{-1}gh$. Then $fS = S$, $fT = T'$ and $fT' = T$. This completes the proof of (2.2).

We complete the proof of the main theorem as follows. We are given that all the oriented edges of G are equivalent under $\mathbf{A}(G)$. Hence there is a greatest positive integer s such that all the s-paths of G are equivalent under $\mathbf{A}(G)$. Assume G is not s-regular.

Consider any $(s+1)$-path $X = (v_0, v_1, \ldots, v_{s-1})$ of G. Let its successors be $Y = (y, v_0, v_1, \ldots, v_s)$ and $Y' = (y', v_0, v_1, \ldots, v_s)$. Let k be one of the elements of $\mathbf{A}(G)$ satisfying $k(v_1, v_2, \ldots, v_{s-1}) = (v_0, v_1, \ldots, v_s)$. Then kv_0 is either y or y', and we may adjust the notation so that $kv_0 = y$. This implies $Y = kX$. By (2.2) there exists $f \in \mathbf{A}(G)$ such that $f(v_0, v_1, \ldots, v_s) = (v_0, v_1, \ldots, v_s)$, $fy = y'$ and $fy' = y$. This implies $Y' = fY = fkX$. Hence X can be transformed into either of its successors by an automorphism of G. Similarly each predecessor of X can be transformed into X, and therefore X can be transformed into either of its predecessors.

Combining these results with (2.1) we see that all the $(s+1)$-paths of G are equivalent under $\mathbf{A}(G)$. But this contradicts the definition of s. The theorem follows.

We should perhaps remark here that in an s-regular cubic graph G the vertices of any s-path are all distinct. For any circuit of G has at least 3 vertices by our definitions and at least $2s - 2$ by **(3, Theorem III)**.

3. 1-regular cubic graphs. Let G be any s-regular graph. Let S_0 be a fixed s-path of G. Any element x of $A(G)$ can be associated with a permutation x' of the set $\mathbf{S}$ of all s-paths of G defined as follows: $x'(hS_0) = hxS_0$ for all $h \in \mathbf{A}(G)$. If $x, y \in \mathbf{A}(G)$ the permutations x', y' satisfy the law $(xy)' = y'x'$. Moreover, the correspondence $x \leftrightarrow x'$ is $1-1$. For $x'(S_0) = xS_0$ for each $x \in \mathbf{A}(G)$, and there is only one element of $\mathbf{A}(G)$ mapping S_0 onto xS_0. It follows that the permutations x' of $\mathbf{S}$ are the elements of a permutation group H of the same order as $\mathbf{A}(G)$.

In **(3)** the symmetry of G is discussed in terms of H. There is a unique pair (r, l) of elements of H mapping S_0 into its two successors. It follows that rS and lS are the two successors of S for each $S \in \mathbf{S}$. Another important element of H, denoted by ξ, reverses the order of the vertices in each s-path S. Thus $\xi^2 = 1$.

It is clear that $\xi r \xi S$ and $\xi l \xi S$ are the two predecessors $r^{-1}S$ and $l^{-1}S$ of S, for each $S \in \mathbf{S}$. There are now two possibilities. In the first alternative $\xi r \xi = r^{-1}$ and $\xi l \xi = l^{-1}$; in the second $\xi r \xi = l^{-1}$ and $\xi l \xi = r^{-1}$. We say that G is of the *first* or *second class* according as the first or second alternative holds. These are the two classes mentioned in the Introduction.

(3.1) *Every 1-regular cubic graph is of the second class.*

Proof. Let G be any 1-regular cubic graph. Let a be any vertex of G. Let its incident edges join it to the three vertices b, c, and d. We may suppose $r(a, b) = (c, a)$ and $l(a, b) = (d, a)$. Then $l^{-1}rl^{-1}(d, a) = l^{-1}r(a, b) = l^{-1}(c, a)$. But $l^{-1}(c, a)$ is distinct from $r^{-1}(c, a) = (a, b)$, and is therefore $(a, d) = \xi(d, a)$. Hence $\xi = l^{-1}rl^{-1}$. Similarly $r^{-1}lr^{-1}(c, a) = \xi(c, a)$ and therefore $\xi = r^{-1}lr^{-1}$. Hence $\xi r\xi = r^{-1}lr^{-1}rl^{-1}rl^{-1} = l^{-1}$. Accordingly G is of the second class.

References

1. H. S. M. Coxeter, *Self-dual configurations and regular graphs*, Bull. Amer. Math. Soc., *56* (1950), 413–455.
2. R. Frucht, *A one-regular graph of degree three*, Can. J. Math., *4* (1952), 240–247.
3. W. T. Tutte, *A family of cubical graphs*, Proc. Camb. Phil. Soc., *43* (1948), 459–474.

University of Toronto

QUADRICS OVER GF(2) AND THEIR RELEVANCE FOR THE CUBIC SURFACE GROUP

W. L. EDGE

1. Introduction. About half the following pages are concerned with the "cubic surface group" G, of order 51840, and the geometry cognate to a certain representation thereof. The literature of this group, with its subgroup of order 25920, is already voluminous; the addition of these few pages to it will not, it is hoped, be regarded askance as a perverse and misdirected indulgence in archaism. On p. 104 of **(2)** one reads

> It is thus suggested that we should interpret the symbols of lines as being symbols of points in space of 5 dimensions, every such point being dependent upon 6 points; the points so arising from the lines of a tritangent plane of the cubic surface will then be upon a line. We are thus led to a configuration of 27 points, in this space, lying in threes upon 45 lines, there being 5 of these lines passing through each of these 27 points.

Precisely such a figure is to hand; one has merely to base the geometry on the finite field $\mathbf{F}$, of two marks 0, 1 with $1 + 1 = 0$. There are, in a projective space [5] of 5 dimensions over $\mathbf{F}$, two projectively distinct types of non-singular quadric; one type has planes on it or, as one may say, is ruled, and the other not. It is this latter quadric that underlies the figure so forcibly suggested to Baker, and the geometry associated with it is summarized in Table II below. This geometry cannot be appreciated properly unless the geometry in the subordinate spaces of lower dimension is known in some detail; the earlier pages aim therefore to assemble this essential information.

To lay these foundations, and so adequately to describe the geometry that underlies this aspect of the cubic surface group, is not, however, the sole purpose of these pages. Once the geometry in [5] is known it may help to open the way towards the geometry, over $\mathbf{F}$, in spaces of higher dimension, of quadrics and their groups of automorphisms. There is a quadric in [7] consisting of 135 points and having on it, in accordance with Study's principle of triality **(8**, p. 477**)**, two systems each of 135 solids; its group of automorphisms, of order $2^{12}3^{5}5^{2}7$, is isomorphic to a group already investigated in a different setting **(7)**; in a [7], yes, but with the complex field, not $\mathbf{F}$, as base field. Just as Table II below is intimately related to Table I in **(7)** so there is a table allied to Miss Hamill's elaborate Table II. The geometry of quadrics over $\mathbf{F}$ is the avenue towards its compilation and any light that it may shed on the structure of the larger group. The table in question was in fact compiled a year or two ago.

Received October 27, 1958.

All the geometry in this paper is based on **F** so that, throughout, a line consists of 3 points, a plane of 7 points lying 3 by 3 on 7 lines, and so on.

2. A symmetric senary quadratic and its automorphisms. A projective space $[n]$, of n dimensions, over **F** consists of $2^{n+1} - 1$ points. Consider, among the 63 points of [5], those whose co-ordinates satisfy

$$\sum_{i,j=0}^{5} x_i x_j = 0, \tag{2.1}$$

the summation being over all pairs i, j with $i < j$, and so embracing 15 products. Such points constitute a quadric $\mathscr{Q}$. Each product $x_i x_j$ is 0 or 1, and a point is on or off $\mathscr{Q}$ according as an even or an odd number of the 15 products is 1. Now if $x_i x_j = 1$ then $x_i = x_j = 1$; hence when a point has

$$1, \quad 2, \quad 3, \quad 4, \quad 5, \quad 6$$

of its 6 co-ordinates non-zero the number of non-zero products is, accordingly,

$$0, \quad 1, \quad 3, \quad 6, \quad 10, \quad 15.$$

Thus $\mathscr{Q}$ consists of

$$\binom{6}{1} + \binom{6}{4} + \binom{6}{5} = 6 + 15 + 6 = 27$$

points. Denote any such point by m, and any of the $63 - 27 = 36$ points that are off $\mathscr{Q}$ by p. This notation will be used throughout to distinguish between points on and points off a quadric.

The partitioning, as $6 + 15 + 6$, of the 27 m is adventitious in that it depends on the system of co-ordinates; but it throws one cardinal feature into strong relief. The first batch of 6 consists of the vertices X_i of the simplex of reference Σ_0; the last batch of 6 consists of the vertices X_i' of another simplex Σ_0'; Σ_0 and Σ_0' are in perspective from the unit point U (**2**, p. 105) and are in the Möbius relation, each being both inscribed and circumscribed to the other. Furthermore, Σ_0 and Σ_0' are both circumscribed to as well as inscribed in $\mathscr{Q}$, the tangent prime of $\mathscr{Q}$ at any vertex of either being a bounding prime of the other. All this is manifest from the equations

$$x_i = 0, \qquad x_i = x_0 + x_1 + x_2 + x_3 + x_4 + x_5 = \sigma,$$

say, of the bounding primes of Σ_0 and Σ_0', and from the co-ordinates of their vertices; all co-ordinates save one of X_i are 0, all save one of X_i' are 1. That Σ_0 and Σ_0' are related to $\mathscr{Q}$ in precisely the same way follows because

$$\sum_{i<j} (\sigma + x_i)(\sigma + x_j) \equiv \sum_{i<j} x_i x_j$$

over **F**; the extra terms on the left amount to $(15 + 5)\sigma^2$ and so, the coefficient being even, disappear.

U is but one of the 36 p. Each p is a centre of perspective for two simplexes,

inscribed and circumscribed to $\mathscr{L}$ and in the Möbius relation. The correspondence with the 36 double-sixes on a cubic surface is unmistakeable.

3. Subject, now, the variables x_i to the linear substitution

$$x = \mathbf{M}\xi$$

where x and ξ are column vectors, each having six components and $\mathbf{M} = (a_{rs})$ is a 6×6 matrix in which each a_{rs} is an element of the field $\mathbf{F}$. The coefficient of ξ_k^2 on the left of 2.1 is $\Sigma_{i<j} a_{ik}a_{jk}$. If the substitution leaves $\mathscr{L}$ invariant each square ξ_k^2 must be absent; hence, foreach k, $\Sigma_{i<j} a_{ik}a_{jk} = 0$; in other words, each column of $\mathbf{M}$ is the co-ordinate vector of a point on $\mathscr{L}$. The column thus includes 1, 4, or 5 non-zero co-ordinates. But it is not enough, to ensure invariance, that each square ξ_k^2 be absent; every product $\xi_i\xi_j$ has to be present. The coefficient of, say, $\xi_1\xi_2$ is $\Sigma_{\alpha\neq\beta} a_{\beta 1}a_{\alpha 2}$, a sum of 30 products, whose vanishing is the condition for the conjugacy of those 2 points of $\mathscr{L}$ whose co-ordinate vectors are the first 2 columns of $\mathbf{M}$. Thus the columns of $\mathbf{M}$ have not only, each of them, to represent points m; it is also necessary that no two of these 6 m be conjugate, that is, that each of their 15 joins has to be a chord c, never a generator g, of $\mathscr{L}$. These conditions, that the columns of $\mathbf{M}$ are co-ordinate vectors of points on $\mathscr{L}$ no two of which are conjugate, are, when supplemented by the proviso that the columns be not linearly dependent, sufficient as well as necessary for the quadratic form to be invariant under the linear substitution.

It might, in building a matrix $\mathbf{M}$ to satisfy these restrictions, be helpful to observe that they imply that each column *is* conjugate to the sum of any two of the others: the conditions

$$\sum_{i\neq j} a_{i\beta}a_{jm} = 1 = \sum_{i\neq j} a_{i\beta}a_{jn}$$

imply

$$\sum_{i\neq j} a_{i\beta}(a_{jm} + a_{jn}) = 1 + 1 = 0.$$

The equation of $\mathscr{L}$ may take other forms than 2.1; see, for one example, § 17 below. Each such form affords, by its group of automorphisms, a representation of G; the matrices that constitute this representation are subject to appropriate restrictions. Some of these forms may well lead to simpler calculations for the order of G, but the symmetric form 2.1 holds the advantage that the restrictions on $\mathbf{M}$ involve all its columns symmetrically.

4. These discussions apply not merely to senary forms but also to forms in other numbers of variables; they afford an opportunity not only of exhibiting, in this context, certain known facts about the cubic surface group but also of investigating other groups. The indispensable prerequisite is a thorough knowledge of the underlying geometry, and this has to be grounded on the

geometry in spaces of lower dimensions. It is advisable to preface the account of this by a section which takes note of the idiosyncrasies due to the base field having characteristic 2, and such a section now follows.

5. The null polarity. A quadratic form

$$f(x) \equiv f(x_0, x_1, \ldots, x_n)$$

is *singular* when it is expressible as a quadratic form in fewer variables x_i', linear forms in the x_i. The quadric $f = 0$ is then a *cone* whose *vertex* V is the $[n - r]$ over which all x_i' simultaneously vanish; r is the rank of the matrix of coefficients in the x_i', and every point of V satisfies $f = 0$. The polar form of any point of V vanishes identically; every point of V is conjugate to every point of $[n]$. The converse is also true *except over fields of characteristic* 2; over such fields a point can be conjugate to every point of $[n]$ without f being singular. If, for example, $n = 2$,

$$f \equiv x_1x_2 + x_2x_0 + x_0x_1$$

is non-singular; yet the point $x_1 = x_2 = x_3$ is conjugate to every point of the plane. This phenomenon can always occur for even n, as is now to be explained.

Recall, first, the genesis of the polar form: it is the coefficient of $\lambda\mu$ in

$$f(\lambda x + \mu y) \equiv f(\lambda x_0 + \mu y_0, \lambda x_1 + \mu y_1, \ldots, \lambda x_n + \mu y_n),$$

and is symmetric in x and y. Since, over $\mathbf{F}$, $\lambda\mu$ has zero coefficient in $(\lambda x_i + \mu y_i)^2$, squares of the variables are ignorable when polarizing; all 2^{n+1} forms

$$a_0x_0^2 + a_1x_1^2 + \ldots + a_nx_n^2 + \sum_{i<j} b_{ij}x_ix_j,$$

with different a_i but fixed b_{ij}, have the same polar $\Sigma b_{ij}(x_iy_j + x_jy_i)$. Whenever $y = x$ every term of this sum vanishes; the polar prime of P *always* passes through P whether P be on $f = 0$ or no; the polarity set up by a quadric is, over $\mathbf{F}$, a null polarity.

6. A point common to the polar primes of the $n + 1$ vertices of the simplex of reference is common to those of all the points of $[n]$, and conversely; such points constitute an $[n - q]$ where q is the rank of

$$\mathbf{B} \equiv \begin{bmatrix} . & b_{01} & b_{02} \ldots b_{0n} \\ b_{01} & . & b_{12} \ldots b_{1n} \\ . & . & \\ . & . & \\ . & . & \\ b_{0n} & b_{1n} & b_{2n} \ldots . \end{bmatrix}$$

there being no common point if $\mathbf{B}$ has its full rank $n + 1$. If n is odd the rank may be full; of this two instances, both convenient to have at hand, may

be noted. First: let $b_{ij} = 0$ except that $b_{i,i+1} = b_{i+1,i} = 1$ for each even i. This matrix answers to the canonical form

$$x_0x_1 + x_2x_3 + \ldots + x_{n-1}x_n = 0$$

of a ruled quadric (that is, of a quadric having on it linear spaces of the maximum possible dimension $\frac{1}{2}(n-1)$), and the non-vanishing of $|\mathbf{B}|$ is obvious. Second: let $b_{ij} = 1$ except for $i = j$. Here too $|\mathbf{B}| = 1$ for odd n, and $\mathbf{B}$ answers to the symmetric form $\Sigma_{i<j} x_ix_j$, with $\frac{1}{2}n(n+1)$ terms in the sum.

7. That $|\mathbf{B}|$ always vanishes when n is even follows because any non-zero term in the determinantal expansion is partnered with another; the expansion of, for example

$$\begin{vmatrix} . & h & g \\ h & . & f \\ g & f & . \end{vmatrix}$$

consists of two equal terms fgh. A set of $n+1$ units in $\mathbf{B}$ whose product supplies a term of $|\mathbf{B}|$ cannot, $n+1$ being odd, be placed symmetrically about the leading diagonal; thus any non-vanishing terms of $|\mathbf{B}|$ are paired, those of a pair being products of sets of $n+1$ units with either set the image of the other in the diagonal. The two products, being equal, have, over $\mathbf{F}$ sum zero. Alternatively: that $|\mathbf{B}| = 0$ for n even is a consequence of $\mathbf{B}$ being skew: over $\mathbf{F}$ a matrix that is symmetric is skew also when its diagonal consists of zeros.

Although the rank q of $\mathbf{B}$ is, when n is even, less than $n+1$, it need not be less than n, for a principal minor can have rank n. Thus, for even n, there may be a single point conjugate to every point of $[n]$. This point need not satisfy $f = 0$. For take the symmetric form

$$f \equiv \sum_{i<j} x_ix_j$$

whose $\mathbf{B}$, when n is even, has $q = n$. The point conjugate to every point of $[n]$ is the unit point

$$x_0 = x_1 = \ldots = x_n = 1,$$

at which $f = \frac{1}{2}n(n+1)$; this is zero or not according as $\frac{1}{2}n$ is even or odd.* And if $\frac{1}{2}n$ is even we only (in order to obtain a non-singular quadric with a point conjugate to every point of $[n]$ and yet not lying on the quadric) need to take for $\mathbf{B}$ the direct sum of $[\begin{smallmatrix}1 & 1\\ 1 & .\end{smallmatrix}]$ and the matrix of the above symmetric form in $n-2$ variables. If, for example, $n = 4$ this leads to the form

$$yz + zx + xy + uv;$$

*When $\frac{1}{2}n$ is even the symmetric form is singular, being the same as $\sum_{0<i<j}(x_0 + x_i)(x_0+x_j)$ wherein x_0^2 occurs an even number, $\frac{1}{2}n(n-1)$, of times and x_0x_i an odd number, $n-1$, of times.

the single point conjugate to every point is $x = y = z = 1, u = v = 0$; it does not lie on the quadric.

8. Whenever a point x on $f = 0$ is conjugate to every point of $[n]$ its join to any other point y on $f = 0$ lies entirely on the quadric because

$$f(\lambda x + \mu y) = \lambda^2 f(x) + 0 + \mu^2 f(y) = 0 + 0 + 0 = 0;$$

this implies that x belongs to the vertex. But its join to any point z not on $f = 0$ does not meet the quadric save at x because

$$f(\lambda x + \mu z) = \lambda^2 f(x) + 0 + \mu^2 f(z) = \mu^2 f(z).$$

Note too that if P, P' are any pair of conjugate points then the line PP' has to meet $f = 0$, if at neither P nor P' than at $P + P'$, because

$$f(x + x') = f(x) + 0 + f(x') = f(x) + f(x').$$

Hence if $q < n - 1$, so that there are at least 2 points conjugate to every point of $[n]$, f is necessarily singular.

9. Suppose then, the base field being $\mathbf{F}$, that n is even and that a non-singular quadric Q is given. There is a unique point k conjugate to every point of $[n]$; it does not lie on Q and may be called its *kernel*. Any point m on Q, being conjugate both to itself and to k, is conjugate also to the third point p on km; p is off Q because the join of two conjugate points on Q is entirely on Q, whereas k is not. Conversely: the third point of the join of k to any point p off Q is on Q because the join of two conjugate points has to meet Q. Thus, of the $2^{n+1} - 2$ points of $[n]$ other than k half are m, on Q, and half are p, off Q; Q consists of $2^n - 1$ points one on each line t_0 through k. The m and p on any t_0 have the same polar prime—the tangent prime of Q at m. There are in $[n]\, 2^n$ primes other than these tangent primes, namely those primes not containing k; these, because of the degeneracy of the null polarity, are not polars of points.

10. Quadrics in [2] and [3]. One now sets out to describe the quadric in the spaces of low dimension over $\mathbf{F}$. The description must cover singular, as well as non-singular, quadrics because singular quadrics occur among the sections of non-singular ones in higher space.

There are 7 binary quadratics over $\mathbf{F}$; of these x^2, $x^2 + y^2$, y^2 are perfect squares, $x^2 + xy$, xy, $xy + y^2$ are reducible and $x^2 + xy + y^2$ is irreducible. The perfect squares are singular; equated to zero they give each a repeated point on the 3-point line L whereon x, y are homogeneous co-ordinates. The other forms are non-singular; the reducible ones give the pairs of points of L, but $x^2 + xy + y^2$ has no zero. These three types of binary quadratic correspond to the three ways in which a line that is not a generator can be related to a quadric Q; there are tangents t, with a single point on Q, chords c, with two points on Q, and lines s that are skew to Q.

A singular quadric in $[n]$ has a vertex $[n-r]$ and projects therefrom a non-singular quadric in $[r-1]$. Hence, taking $n=2$, singular conics are of three types; repeated lines ($r=1$) and either line-pairs or single points ($r=2$). A non-singular conic consists of 3 points A, B, C, one on each line through its kernel k; of the 7 lines in its plane, 3, namely BC, CA, AB, are c, 3, namely kA, kB, kC, are t, while the remaining line s is skew to the conic. And a plane section of a quadric Q in higher space may be by a plane

e, meeting Q in a single line,
f, meeting Q in two lines,
h, meeting Q in a single point,
j, meeting Q in 3 non-collinear points.

It may also happen, if $n \geqslant 5$, that a plane d lies wholly on Q. But no plane can ever be skew to Q. Since a singular quadric in [3] is the projection of a conic from a point, one can now tabulate the following sections of Q; a section may be by a solid

γ, meeting Q in a single plane,
ϕ, meeting Q in two planes,
χ, meeting Q in a single line,
ψ, meeting Q in 3 concurrent non-coplanar lines.

It may also happen, if $n \geqslant 7$, that a solid ω lies wholly on Q. There are also solids κ meeting Q in ruled quadrics, and solids λ, meeting Q in non-ruled quadrics; these sections are not singular, and will be described in a moment. The letters here used to signify points, lines, planes, solids in their various relations to Q will be retained throughout. The nature of the subspaces in a given space is recognizable at a glance. In, say χ, each of the 3 planes through the g is an e while the other 12 planes are all h; the 16 lines skew to g are all s while the 18 which, 6 at each of its 3 points, meet g are all t. The points in an e other than the 3 m on g are a quadrangle of p; these quadrangles, like a triad of desmic tetrahedra, are such that any two of them are in perspective from every vertex of the third. Geometry in solids other than χ is clear too; there will be occasion to speak of ψ in § 18 below.

11. A non-singular ruled quadric, or hyperboloid, $\mathscr{H}$ in [3] has the canonical equation $xy+zt=0$; there are on $\mathscr{H}$ two complementary reguli with 3 g in each. Thus $\mathscr{H}$ consists of 9 m; at each m the tangent plane f is spanned by lines one in each regulus; the third line t in f through m does not meet $\mathscr{H}$ again and is, like each g, its own polar. The 6 points of [3] off $\mathscr{H}$ lie (4, p. 323) 3 on each of two lines s, s'; these are polar lines with respect to $\mathscr{H}$. They may be called the Dandelin lines of $\mathscr{H}$, for Dandelin's classical construction sets out from 3 lines in each of two complementary reguli and is viable over any field. When the field is $\mathbf{F}$ the construction utilizes every point of the space.

12. That the quadric $\mathfrak{p}$, whose equation is

$$xy + xz + xt + yz + yt + zt = 0, \tag{12.1}$$

is non-singular was implied in § 6. The vertices of the tetrahedron of reference T all lie on $\mathfrak{p}$, and each face of T meets $\mathfrak{p}$ in the three vertices of T therein and nowhere else; should any co-ordinate be zero no point other than the vertices of T satisfies 12.1. But there is one point, namely the unit point, none of whose co-ordinates is zero, and this does satisfy 12.1. $\mathfrak{p}$ is, indeed, simply a pentagon: a set of 5 points no 4 of which are coplanar. There would perhaps be no call to treat a pentagon as a quadric if spaces of higher dimensions were not to be encountered; but these pentagons are solid sections of quadrics in higher space and their properties call for interpretation in this context. The pentagon has, for instance, tangent planes; these are

$$y + z + t = 0, \qquad z + t + x = 0, \qquad t + x + y = 0, \qquad x + y + z = 0$$

at the vertices of T, and $x + y + z + t = 0$ at the unit point. They are planes h; whereas the 10 faces of the pentagon are all j. Each j has for its pole the kernel of the conic that is its intersection with $\mathfrak{p}$; this kernel is the point common to j and the join c of the two points of $\mathfrak{p}$ outside j. The 35 lines of [3] are distributed as 10 c (edges of $\mathfrak{p}$), 10 s (1 in each j) and 15 t (3 through each vertex of $\mathfrak{p}$ and lying in the tangent plane h there). Each t is its own polar, but c and s are polars of each other.

13. The non-singular quadric in [4]. A non-singular quadric Q in [4] consists of 15 m, one on each of the 15 lines t_0 through its kernel k. The tangent solid at m contains k, and through k there pass, in this solid, 6 lines other than km. But each such line contains a point of Q which, being conjugate to m, is joined to m by a g; thus each of the three planes which lie in the solid and pass through km contains a g through m, and these g form the section of Q by the solid, a solid which may therefore be designated by ψ. Q is thus the exemplar, over $\mathbf{F}$, of the figure (**1**, frontispiece) of 15 points lying 3 by 3 on 15 lines in [4]. The 3 g through an m span ψ, and one idiosyncrasy of the figure consequent upon the base field being $\mathbf{F}$ is that the 15 ψ are concurrent (at k).

Let m_1 be a given m. There are $15 - 1 - 6 = 8$ m not conjugate to m_1 and whose joins to m_1 are therefore c; hence Q has $\frac{1}{2}(15 \times 8) = 60$ c. The plane ck contains not only m_1, m_2, say on c but also m_3 on the join of k to that point p on c other than m_1, m_2 themselves. Hence there are 20 planes j_0 through k, each meeting Q in 3 m that form a triangle whose sides are all c; such a plane includes a line s skew to Q. As there are in all 35 planes through k they are accounted for by the 20 j_0 and by the 15 e which join k to the 15 g.

The polar of a line μ not passing through k is a plane through k and lies in the polar solids of all points in the plane μk. If μ is g the polar is gk. But if μ is s the polar will be j_0', skew to s because no two points on s are conju-

gate. Thus the 20 s are paired; the joins j_0, j_0' of a pair s, s' to k are polars of s', s respectively. Every point on s is conjugate to every point on s'; the 9 transversals of s, s' are all t and the solid κ spanned by s and s' meets Q in a hyperboloid with s, s' for its Dandelin lines. These 10 solids $\kappa = [s\, s']$, containing 6 g that fall into complementary reguli of 3 lines each, are the exemplars of what Baker **(1**, p. 115**)** calls the singular solids. When lines are skew in Baker's frontispiece the corresponding lines on Q are skew; hence the 15 g form 6 quintuples of mutually skew lines, each g belonging to 2 quintuples and being determined as their common line. Each κ answers to a partition of the 6 quintuples into complementary triads and it is, as in the classical case, a consequence of this that the plane common to two solids κ_1, κ_2 is a plane f of a pair of intersecting g. The third solid through f is, of course, the ψ which joins it to k.

Of the solids in [4] 16 do not contain k; those 6 of these that are not κ are λ; each contains 5 m and the triads of g concurrent thereat together account for all 15 g. In the analogous classical figure the corresponding pentagons certainly occur: each of the 15 points corresponds, via the 3 lines concurrent there, to a syntheme (that is, a partitioning into three pairs)of the 6 quintuples; when 5 synthemes constitute a synthematic total(that is, they together account, by the 3 pairs in each, for all 15 pairs of quintuples) the 5 corresponding points form one of the pentagons in question. Since there are 6 synthematic totals there are 6 pentagons; but the vertices of a pentagon need not lie in a solid. That they do so here is another consequence of the base field having characteristic 2.

14. An incidence table for Q is now given; a zero suffix indicates that a space passes through k, whereas its counterpart without the suffix does not. The upper half, above the diagonal, shows the number of subspaces, indicated by a letter on the extreme right, that lie in a given space indicated by a letter at the top. The bottom half, below the diagonal, shows the number of spaces, again indicated by a letter on the right, that contain a given subspace. Either half is deducible from the other: for example, since there are 18 of the 60 c in each of the 10 κ the number of κ through a given c is $18 \times 10/60 = 3$.

15. The whole figure can be described succinctly by using a supernumerary set of 6 homogeneous co-ordinates whose sum is zero. Then the 15 points having two zero co-ordinates constitute Q, and collinearities occur for points such as

$$(0, 0, 1, 1, 1, 1); \qquad (1, 1, 0, 0, 1, 1); \qquad (1, 1, 1, 1, 0, 0).$$

Thus in this representation it is the g, not the m, that correspond to synthemes; the m, not the g, that correspond to pairs. Q is the section of the non-singular quadric $\Sigma_{i<j} x_i x_j = 0$ in [5] by the [4]

$$x_0 + x_1 + x_2 + x_3 + x_4 + x_5 = 0.$$

TABLE I

1	15	15	15	15	45	60	20	15	15	20	60	45	15	10	6	
k	p	m	g	l_0	l	c	s	h	e	j_0	j	f	ψ	κ	λ	
			.	1	.	.	.	.	1	1	.	.	1	.	.	k
			.	1	2	1	3	6	3	3	4	2	7	6	10	p
			3	1	1	2	.	1	3	3	3	5	7	9	5	m
.	.	3						.	1	.	.	2	3	6	.	g
15	1	1						.	3	3	.	.	7	.	.	l_0
.	6	3						3	3	.	3	1	9	9	15	l
.	4	8						.	.	3	3	4	12	18	10	c
.	4	.						4	.	1	1	.	4	2	10	s
.	6	1	.	.	1	.	3						1	.	5	k
15	3	3	1	3	1	.	.						3	.	.	e
20	4	4	.	4	.	1	1						4	.	.	j_0
.	16	12	.	.	4	3	3						4	6	10	j
.	6	15	6	.	1	3	.						3	9	.	f
15	7	7	3	7	3	3	3	1	3	3	1	1				ψ
.	4	6	4	.	2	3	1	.	.	.	1	2				κ
.	4	2	.	.	2	1	3	2	.	.	1	.				λ

Examples of equations for subspaces of this [4] are as follows.

$$\lambda : x_1 = 0.$$
$$\kappa : x_1 + x_2 + x_3 = 0 = x_4 + x_5 + x_6.$$
$$\psi : x_1 + x_2 = 0.$$
$$h : x_1 = x_2 = 0.$$
$$f : x_1 + x_2 = x_3 = x_4.$$
$$j_0 : x_1 = x_2 = x_3.$$
$$j : x_1 + x_2 = x_3 = 0.$$
$$e : x_1 + x_2 = x_3 + x_4 = x_5 + x_6.$$
$$s : x_1 = x_2 = x_3 = 0.$$
$$c : x_1 = 0,\ x_2 = x_3 = x_4.$$
$$l : x_1 = x_2 = x_3 + x_4 = x_5 + x_6.$$
$$l_0 : x_2 = x_3 = x_4 = x_5.$$

The 6 homogeneous co-ordinates related by $\Sigma x_i \equiv 0$ were used, though not over a finite field, by Castelnuovo (**3**, p. 885) in his study of linear complexes in [4]. Such a complex sets up, as does Q, a null polarity which, as with Q, is degenerate (which it cannot but be in a space whose dimension is even); the kernel of Q is analogous to what Castelnuovo calls the centre of the complex. Castelnuovo's statement (**3**, p. 862) that it is from its centre, and its centre only, that the lines of a linear complex in [4] are projected into those of one in [3], is valid in the finite figure. For, since the lines of a screw through a point of [3] are coplanar, any eligible centre of projection has to lie in the

solid that contains the lines of the complex at any point of [4]; but the g at a point of Q span ψ, and the only point common to all ψ is k. Thus projection from k onto a [3] which must, as not passing through k, be κ or λ, gives 15 points lying 3 by 3 on 15 lines, the 3 lines through any point being coplanar. These lines constitute a screw **(4, p. 325)**. If the projection is onto κ it is at once seen that the 15 lines consist of triads forming complementary reguli and of the tangents to the quadric, one at each of its 9 points; that these constitute a screw is known. (The map of complementary reguli on the Klein quadric in [5] is by conics in a pair of polar planes: when the base field is F these planes intersect at the kernel of both conics, and so lie in a [4].) This accords with the known isomorphism **(9, p. 17)**, when the base field has characteristic 2, of the orthogonal group on 5 variables with the symplectic group on 4.

16. The non-ruled quadric in [5]. It is now, with the facts assembled in the preceding sections, a straightforward matter to give a conspectus in [5] of the geometry of $\mathscr{L}$, the quadric introduced in § 2. $\mathscr{L}$ consists of 27 m; the tangent prime M at m meets $\mathscr{L}$ in a cone that projects a $\mathfrak{p}$ from a point outside its [3]; instances are the bounding primes of Σ_0 and Σ_0' (cf. the footnote to § 7). The polar prime P of any of the 36 p off $\mathscr{L}$ meets $\mathscr{L}$ in a non-singular quadric having p for kernel; the 15 lines through p in P are all t, and are the only t through p; of the 16 lines through p outside P 6 are (§ 2) c; the remaining 10 must be s. Since the numbers of p on t, c, s are 2, 1, 3 the total number in [5]

$$\begin{aligned} \text{of } t \text{ is } 36 \times 15/2 &= 270, \\ \text{of } c \text{ is } 36 \times 6 &= 216, \\ \text{of } s \text{ is } 36 \times 10/3 &= 120. \end{aligned}$$

Any other line is a g, and as there are, through any m, 5 g projecting the vertices of a $\mathfrak{p}$ the total number of g on $\mathscr{L}$ is $27 \times 5/3 = 45$.

Each line has a polar solid. The polar of c is λ: for example, each bounding [3] of either Σ_0 or Σ_0' is the polar of an edge of the other. The polar of s is κ: for example, $x_4 = x_5 = x_0$ is the polar of $x_1 = x_2 = x_3 = x_4 + x_5 + x_0 = 0$. As for the polars of t and g, let m_0 be any point of $\mathscr{L}$. All t and all g through m_0 lie in the tangent prime M_0; hence their polar solids all lie in M_0 and contain m_0. They thus project, from m_0, the polar planes j, h with respect to $\mathfrak{p}$ of points p, m in the [3] wherein $\mathfrak{p}$ lies; thus t has a polar solid ψ and g a polar solid χ.

It only remains, before compiling Table II, to discuss the planes in [5] and their relations to $\mathscr{L}$. A plane through g in the polar χ of g meets $\mathscr{L}$ in g only, and so is an e; e is its own polar, every point therein being conjugate to every other. There are 135 e, 3 through each g. Also there are 270 f, spanned by pairs of g. Since f is spanned by a pair of g its polar is the intersection of a pair of χ and so is an h—the only type of plane other than e to lie in χ. The

10 f and 10 h through m project the 10 c and 10 s of a $\mathfrak{p}$. Finally there are j, for instance the plane faces of Σ_0. Since j contains an s its polar lies in a κ and so is also a j—the only type of plane other than f to lie in κ. Now while each κ contains 6 j the κ through a given j is, with the s in the polar j, unique. Hence there are 720 j, consisting of 360 polar pairs. The kernel of the section of $\mathscr{L}$ by j, being conjugate to every point of j, lies in the polar plane; thus the j of a polar pair have in common the kernel of both their sections of $\mathscr{L}$. There are 10 of the 360 polar pairs through each of the 36 p, and they are (§ 13) the pairs of polar j_0 in P.

17. Table II can now be compiled. In order that it be endowed with central symmetry one refers the lower half, registering the numbers of spaces through any given subspace, to the column of symbols on the left and the row of symbols below; the upper half, registering the numbers of subspaces in any given space, is indexed, as was Table I, by means of the row above and the column on the right. This enables one to arrange the symbols on the left and right so that those in either column indicate the polar spaces of those in the other read in *exactly the reverse order*; the same applies to the rows at the top and bottom.

A notable feature of the geometry is the distribution of the 120 s in trios s, s', s'' such that the polar κ of each member of the trio is the solid spanned by the other two. If the vertices of the simplex of reference were chosen 2 on each of s, s', s'' the equation of $\mathscr{L}$ would be, since the quadratic giving the intersections with any of s, s', s'' is irreducible,

$$y_0^2 + y_0y_1 + y_1^2 + y_2^2 + y_2y_3 + y_3^2 + y_4^2 + y_4y_5 + y_5^2 = 0.$$

G is thus representable as the group of automorphisms of this quadratic form, and its order is instantly calculable now it is known that there are 40 trios. For there are on each member of a trio 3 choices for the pair of reference points which, having been chosen, may yet be transposed; also the members of the trio may undergo permutation. Hence G has order

$$40 . (3.2)^3 3! = 51840.$$

18. Table II is intimately related to Table I of **(7)**; it is natural that this should happen because the two tables depict closely allied representations of the same group—the cubic surface group of order 51840. Both representations are in projective space of 5 dimensions, but here the base field **F** is finite whereas in **(7)** it is the field of complex numbers. The main feature of contrast, among a multitude of similarities, is, however, that no analogues of the 27 m appear in **(7)**, where the whole figure is based on 36 points p_0; this forestalls any possibility of there appearing in **(7)** the analogues of any spaces whose p are inadequate to span them; that is, of g, whereon no p occurs, of c, whereon there is but a single p, of f, whereon the p are collinear.

TABLE II

	36 p	27 m	45 g	270 t	216 c	120 s	270 h	135 e	720 j	270 f	120 κ	216 λ	270 ψ	45 χ	27 M	36 P	
p			.	2	1	3	6	4	4	2	6	10	8	12	20	16	p
m			3	1	2	.	1	3	3	1+4	9	5	1+6	3	1+10	15	m
g	.	5					.	1	.	2	6	.	3	1	5	15	g
t	15	10					3	6	3	1	9	15	4+12	18	10+60	60	t
c	6	16					.	.	3	4	18	10	12	.	40	60	c
s	10	.					4	.	1	.	2	10	4	16	40	20	s
h	45	10	.	3	.	9					.	5	1	12	10+40	15	h
e	15	15	3	3	.	.					6	10	8	.	80	80	j
j	80	80	.	8	10	6					.	.	3	3	15	15	e
f	15	10+40	12	1	5	.					9	.	3	.	10	45	f
κ	20	40	16	4	10	2	.	1	.	4					.	10	κ
λ	60	40	.	12	10	18	4	3	.	.					16	6	λ
ψ	60	10+60	18	4+12	15	9	1	3	6	3					10	15	ψ
χ	15	5	1	3	.	6	2	.	1	.					5	.	χ
M	15	1+10	3	1+6	5	9	1+4	3	3	1	.	2	1	3			M
P	16	20	12	8	10	6	2	4	4	6	3	1	2	.			P
	p	m	g	t	c	s	h	j	e	f	κ	λ	ψ	χ	M	P	

Certain entries are shown as sums of two numbers. This indicates, in the upper right-hand half of the table, that when a space meets $\mathscr{S}$ in a singular quadric subspaces of the same dimension may be related in different ways to the vertex of the section; and dually for the lower left-hand half of the table. For example: of the 16 t in ψ, 4 pass through the concurrence of the 3 g therein, whereas the other 12 do not; and of the 50 h in an M, 10 pass through the contact m, whereas the other 40 do not.

Miss Hamill's nomenclature for the subspaces of the figure in (7) is easily translated into that used here, but she builds solids and primes of more varied kinds and in larger numbers, and one must explain why this is. The following short dictionary

p_0	e_1	γ_1	$\gamma_1 \times p_0$	c_2	$\gamma_1 \times \gamma_1$
p	t	s	j	h	κ

shows how spaces whose symbols are below are denoted, in (7), by the symbols above: e_1 is defined as a line spanned by two conjugate p_0 whereas t is spanned by two conjugate p, while $\gamma_1 \times \gamma_1$ is the translation of the fact that κ is spanned by its Dandelin lines. But consider, for a moment, a solid ψ. It contains g_1, g_2, g_3, concurrent at m_0, and 8 j, namely those planes in ψ which do not contain m_0, consisting of 4 pairs, a pair together containing all points on g_1, g_2, g_3 other than m_0. The intersection of such a pair of j is an s, and the 4 s are joined to m_0 by the same h—the only plane in ψ through m_0 that does not contain any of g_1, g_2, g_3. Now the kernel of the conic in any of the 8 j is on t, the line in ψ through m_0 that is not in any of the planes g_2g_3, g_3g_1, g_1g_2; thus, m_0 itself being conjugate to each point in ψ, each point of t is conjugate to each point of h. Hence ψ may be spanned by t and any of the s in h, and it is because there are 4 choices for this s that there occur 1080 solids $\gamma_1 \times e_1$, in (7)—4 times the number 270 of ψ. Moreover ψ can be spanned by h and either of the p on t, and it is because there are 2 choices for this p that there occur 540 solids $c_2 \times p_0$ in (7). Likewise for primes. There are in P 6 λ, which explains why there are 6 times as many primes $\gamma_2 \times p_0$ in (7) as there are P; and there are 20 s falling into 10 pairs, each pair spanning a κ as its Dandelin lines, which explains why there are 10 times as many primes $\gamma_1 \times \gamma_1 \times p_0$ as there are P.

19. The cubic surface group. The order of the group of senary linear transformations that leave $\mathscr{L}$ invariant can be calculated by appealing to the properties (§ 3) that the columns of the corresponding matrices **M** must have. One has to choose a hexad of m whose 15 joins are all c. For m_1 there are 27 choices; then 16, the number of c through m_1, for m_2; then 10, the number of j through c, for m_3; m_4 has to lie in a λ through j, and there are 3 such λ, but each contains two further m, vertices of a $\mathfrak{p}$ in λ, that are not conjugate to any of m_1, m_2, m_3. Thus there are 6 choices for m_4. Denote, for the moment, the vertex of $\mathfrak{p}$ that is not chosen by m'; it is not then eligible to furnish another column of **M**, since there must be no linear dependence between the 6 columns. On each of

$$m_2m_3, \quad m_3m_1, \quad m_1m_2, \quad m_1m_4, \quad m_2m_4, \quad m_3m_4$$

there is a p, and this has to be (see the remark towards the end of § 3) conjugate to each of the m yet to be chosen. The six p are coplanar, constituting (§ 12) with m' the tangent plane to $\mathfrak{p}$ at m'; their plane h has for its polar an

f wherein the 2 g meet at m'. The polar c of λ lies in f, and the m thereon are not eligible; the others, one on each g, are, and complete the tally. Hence the number of $\mathbf{M}$ is

$$27.16.10.6.2 = 51840.$$

It is true that the classical properties of the 27 lines on a cubic surface also afford a means of obtaining this number, and have long done so; the choice for m_4 after m_1, m_2, m_3 are fixed is analogous to the choice, in Schläfli's notation, of say, a_4 or c_{56} towards filling half a double-six once a_1, a_2, a_3 are fixed. But for corresponding "orthogonal" groups in other numbers of variables there is no such precedent.

20. The 6-rowed permutation matrices π constitute, 2.1 being symmetric in all 6 co-ordinates, a subgroup $\mathscr{S}_6$ of G. Each π has certain latent column vectors; the nature of the space σ of whose points these are co-ordinate vectors depends on the cyclic decomposition of the permutation, and so on a partition of 6. For example: the partition 3^2 corresponds to the cyclic decomposition $(xxx)(xxx)$ and so to a σ with equations of the type

$$x_0 = x_1 = x_2, \qquad x_3 = x_4 = x_5;$$

this is a line s. For a partition $\lambda_1\lambda_2 \ldots$, σ has dimension $5 - \Sigma(\lambda_i - 1)$. No two among the 11 partitions yield the same type of space for σ, so that the π provide representatives of 11 conjugate sets in G.

21. If π is given, one can find the number of matrices conjugate thereto in G by calculating the order of the normalizer. One writes down the most general matrix, with all its elements in $\mathbf{F}$, that commutes with π, and then imposes all the conditions on its columns that are necessary for it to belong to G. If N is the number of such matrices the number of operations in G that are conjugate to π is $51840/N$. As an example take

$$\pi = \begin{bmatrix} . & 1 & . \\ . & . & 1 \\ 1 & . & . \end{bmatrix} \oplus \begin{bmatrix} . & 1 & . \\ . & . & 1 \\ 1 & . & . \end{bmatrix}$$

answering to 3^2. Any matrix that commutes with π has the form

21.1
$$\begin{bmatrix} a_1 & a_2 & a_3 & c_1 & c_2 & c_3 \\ a_3 & a_1 & a_2 & c_3 & c_1 & c_2 \\ a_2 & a_3 & a_1 & c_2 & c_3 & c_1 \\ b_1 & b_2 & b_3 & d_1 & d_2 & d_3 \\ b_3 & b_1 & b_2 & d_3 & d_1 & d_2 \\ b_2 & b_3 & b_1 & d_2 & d_3 & d_1 \end{bmatrix}.$$

If $a_1 + a_2 + a_3 = b_1 + b_2 + b_3 = 0$, this is singular. Thus

either $$a_1 + a_2 + a_3 = b_1 + b_2 + b_3 = 1,$$

when it is necessary, for 21.1 to belong to G, that

$$a_2a_3 + a_3a_1 + a_1a_2 + b_2b_3 + b_3b_1 + b_1b_2 = 1,$$

or one of $a_1 + a_2 + a_3$ and $b_1 + b_2 + b_3$ is 0 and the other 1, when

$$a_2a_3 + a_3a_1 + a_1a_2 = b_2b_3 + b_3b_1 + b_1b_2 = 1.$$

It is not possible to satisfy simultaneously the two conditions

$$yz + zx + xy = x + y + z = 0;$$

they represent a conic and the line in its plane that is skew to it. This enables one to tabulate the whole set of solutions of the above conditions. The outcome is, since c, d are subject to precisely the same restrictions as a, b, that the first 3, as the last 3, columns of 21.1 must be one of the blocks

1	.	.	1	1	1	.	1	1	1	1	1
.	1	.	1	1	1	1	.	1	1	1	1
.	.	1	1	1	1	1	1	.	1	1	1
1	1	1	1	.	.	1	1	1	.	1	1
1	1	1	.	1	.	1	1	1	1	.	1
1	1	1	.	.	1	1	1	1	1	1	.

or else be derived therefrom by cyclic permutation of its 3 columns. The order of the normalizer of π is therefore, since the same block must not occur twice, $4.3.3^2 = 108$, and π belongs to a conjugate set of 480 matrices in G.

The above 12 columns are co-ordinate vectors, of points outside $x_0 + x_1 + x_2 = x_3 + x_4 + x_5 = 0$, no two of which are conjugate.

22. There is another way of finding how many operations compose a conjugate set in G once the space σ constituted by the invariant points is known. One first writes down the most general matrix, with all its elements in $\mathbf{F}$, having all the points of σ among its latent column vectors: this is secured by stipulating the invariance of a set of points that span σ. One next imposes the restrictions on the columns of this matrix to ensure that it belongs to G. One has then to exclude any matrices that leave invariant not merely the points of σ itself but also those of some space of higher dimension that contains σ; this last stage is liable to be the more complicated the lower the dimension of σ. As an example, let σ be a plane h. This does not happen for the permutation matrices, but it illustrates the points at issue equally well and provides another conjugate set. Take, for h,

$$x_0 = x_1 = x_2 + x_3 + x_4 + x_5 = 0.$$

Every point therein is invariant if 3 linearly independent points are; hence one requires a matrix having

$$\begin{matrix} . & . & . \\ . & . & . \\ 1 & . & . \\ 1 & 1 & . \\ . & 1 & 1 \\ . & . & 1 \end{matrix}$$

among its latent column vectors, and so having its last 4 columns of the form

$$\begin{matrix} a & a & a & a \\ b & b & b & b \\ c+1 & c & c & c \\ d & d+1 & d & d \\ e & e & e+1 & e \\ f & f & f & f+1 \end{matrix}$$

Since each, as the co-ordinate vector of a point, must be conjugate to the sum of any two others, $c = d = e = f$; and since each must be the co-ordinate vector of an m,

22.1 $$(a+1)(b+1) = c+1.$$

One now has to prefix 2 columns, each conjugate to the sum of any 2 of the last 4; thus there ensues

$$\begin{bmatrix} A_1 & A_2 & a & a & a & a \\ B_1 & B_2 & b & b & b & b \\ \alpha & \beta & c+1 & c & c & c \\ \alpha & \beta & c & c+1 & c & c \\ \alpha & \beta & c & c & c+1 & c \\ \alpha & \beta & c & c & c & c+1 \end{bmatrix},$$

The conditions on the first two columns are

$$A_1B_2 + A_2B_1 = 1, \qquad A_1B_1 = A_2B_2 = 0.$$

These allow $A_1 = B_2 = 1$, $A_2 = B_1 = 0$; but this is inadmissible since it permits every point of the space $x_0 + x_1 = x_2 + x_3 + x_4 + x_5 = 0$ to be invariant. Hence

$$A_1 = B_2 = 0, \qquad A_2 = B_1 = 1.$$

Next, the first and third columns would be conjugate unless $\alpha = a$, the second and third unless $\beta = b$; moreover α, β must be unequal, for otherwise

the $\psi\ x_0 + x_1 = x_2 + x_3 + x_4 + x_5 = 0$ would have all its points invariant. Since a, b are thus unequal, 22.1 gives $c = 1$, and the matrix becomes

$$\begin{bmatrix} . & 1 & a & a & a & a \\ 1 & . & a+1 & a+1 & a+1 & a+1 \\ a & a+1 & . & 1 & 1 & 1 \\ a & a+1 & 1 & . & 1 & 1 \\ a & a+1 & 1 & 1 & . & 1 \\ a & a+1 & 1 & 1 & 1 & . \end{bmatrix}.$$

Hence there are, corresponding to $a = 0$ and $a = 1$, two matrices in G leaving invariant the points of a given h and no point outside this plane. That they are conjugate in G follows because each is obtainable from the other by simultaneous transpositions of the first two rows and of the first two columns, and so by transformation by a permutation matrix. Each is inverse to the other, and both have for their square an involution for which every point in the $\psi\ x_0 + x_1 = x_2 + x_3 + x_4 + x_5 = 0$ is invariant. They are representative of a conjugate set of 540 operations of period 4 whose invariant points consist of 6 p and an m, a fact that must be mirrored in other representations of G.

23. The following list shows, for each partition of 6, the type of equations for the space σ of invariant points of the corresponding permutation matrices; it also says what kind of space σ is, gives the Roman numeral that labels the corresponding conjugate set in **(6)** and **(5)**, and the number of operations of G in this set. Several facts that are evident from this list have to agree with known results; the period of an operation can be seen from the partition, as can the nature of the different powers. For instance: since 6 has 3^2 for its square and 2^3 for its cube, the operations labelled XXIII must have their squares labelled IX and their cubes labelled XVII. Moreover, the numbers of m and p that are invariant for a given operation, being those of points in its invariant space, are known and must be the same as numbers already encountered in other representations of G.

TABLE III

1^6		[5]	I	1
$1^4 2$	$x_1 = x_2$	P	XVI	36
$1^3 3$	$x_1 = x_2 = x_3$	κ	VI	240
$1^2 4$	$x_1 = x_2 = x_3 = x_4$	f	XVIII	1620
$1^2 2^2$	$x_1 = x_2, x_3 = x_4$	ψ	II	270
123	$x_1 = x_2, x_3 = x_4 = x_5$	j	XXI	1440
15	$x_1 = x_2 = x_3 = x_4 = x_5$	c	XV	5184
6	$x_0 = x_1 = x_2 = x_3 = x_4 = x_5$	p	XXIII	4320
24	$x_0 = x_1, x_2 = x_3 = x_4 = x_5$	l	V	3240
2^3	$x_0 = x_1, x_2 = x_3, x_4 = x_5$	e	XVII	540
3^2	$x_0 = x_1 = x_2, x_3 = x_4 = x_5$	s	IX	480

24. Not every type of subspace in [5] has appeared in the role of σ. Some cannot. Were λ invariant so would its polar c be, and on c lies a single p which would be invariant too. Thus the hypothesis that every point of λ is invariant, implying the invariance of a p outside λ, implies the invariance of every point of the join P of λ and p. Nor, as is easily seen by taking a face of Σ_0, can every point of an M be invariant. It has, however, been seen that σ might be h; this gives the set XIX. Furthermore, every point of

$$\chi : x_0 + x_1 = x_2 + x_3 = x_4 + x_5$$

is invariant for the involution

$$\begin{bmatrix} 1 & . & 1 & 1 & 1 & 1 \\ . & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & . & 1 & 1 \\ 1 & 1 & . & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & . \\ 1 & 1 & 1 & 1 & . & 1 \end{bmatrix}$$

and for no other operation of G save I. This involution is one of 45 in the set III.

25. The 36 involutions J associated one with each p are conspicuous as being the only operations in G whose invariant points fill a [4]. Suppose that p_1, p_2 are such that $J_1J_2 = J_2J_1$. If m is outside both P_1 and P_2, and if m_1 and m_2, the remaining points on mp_1 and mp_2 respectively, are those points into which m is transformed by J_1 and J_2, commutation demands that the intersection m_{12} of m_1p_2 and m_2p_1 is on $\mathscr{L}$. The quadrangle $mm_1m_2m_{12}$ has p_1, p_2 for diagonal points; since **F** has characteristic 2 the third diagonal point is the remaining point of p_1p_2. Now any plane that includes 4 m has to be an f and include 5; hence p_1p_2 is a t. Conversely, let p_1p_2 be a t, ψ its polar solid; as p_1, p_2 are conjugate P_1 and P_2 both contain ψ. Every point in ψ is invariant for both J_1 and J_2; the 16 points of P_1 outside ψ, collinear in pairs with p_2, are transposed in pairs by J_2; the 16 points of P_2 outside ψ are transposed in pairs by J_1; hence the successive application of J_1, J_2 has the same effect, in either order, on any point in either P_1 or P_2. Any point of [5] that is outside both P_1 and P_2 is on lines that meet them both, and is the only point on any such line, it follows that it, too, is changed into the same point by successive application of J_1, J_2, in either order. Thus $J_1J_2 = J_2J_1$.

This discussion has shown that J_1, J_2 do or do not commute according as p_1p_2 is a t or an s. Thus non-commuting J occur (**6**, p. 86) in associated sets of 3. The product of any number of mutually commutative J is itself an involution; each factor is associated with a p and every pair among these p has its join a t. In particular: the 4 points in an e not on the g therein have this property; the product of the 4 corresponding J is the involution whose

invariant points compose the polar χ of g. As there are 3 e through g in χ there are **(6**, p. 88**)** three ways of expressing this same involution as a product of 4 commutative J.

The conjugate sets in Table III are identified by representatives that impose certain permutations on the vertices of Σ_0. The mutually disjoint cycles that compose the permutation correspond, when appropriately expressed as products of transpositions, to those r-chains used by Frame that have $r \leqslant 6$. This is why the same 11 partition labels have appeared among the 18 at the bottom of p. 91 of **(6)**.

26. G is a subgroup, of index 28, of the "group of the bitangents," and this larger group happens too to have a representation in 6 variables over $\mathbf{F}$, being isomorphic to the symplectic group of matrices μ for which $\mu'\mathbf{S}\mu = \mathbf{S}$, when $\mathbf{S}$ is a non-singular skew matrix. A referee has proposed that this occurrence as a subgroup of the symplectic group be shown in the present context, and so this § 26 is appended.

Any projectivity that leaves $\mathscr{L}$ invariant turns conjugate points into conjugate points and non-conjugate points into non-conjugate points; hence, with $\mathscr{L}$ in the symmetric form of § 2, the bilinear form $\sum_{i,j} y_i z_j$ must be unaltered under each of the 51840 projectivities. But this bilinear form is $y'\mathbf{B}z$, z being the column vector whose 6 components are the z_j, y' the row vector with components y_i, and $\mathbf{B}$ the non-singular 6-rowed matrix mentioned at the end of § 6. But $\mathbf{B}$, symmetric and with a zero diagonal, is skew over $\mathbf{F}$, so that all 51840 $\mathbf{M}$ belong to the symplectic group Γ on $\mathbf{B}$.

The condition that $\mathbf{M}$ belong to Γ is less stringent than the condition that $\mathbf{M}$ belong to G; for the symplectic property is not that a quadric but that a polarity is left invariant and, as was noted in § 5, the same polarity arises from 64 quadrics. The wider latitude of the symplectic condition permits these quadrics to undergo permutation. Yet a ruled quadric can be permuted only with other ruled quadrics, and likewise for quadrics, such as $\mathscr{L}$, that are not ruled. One presumes, knowing the index of G in Γ, that 28 of the 64 quadrics are not ruled while the remaining 36 are ruled.

All 64 quadrics, inducing as they do a non-degenerate polarity, are non-singular; in order that one be ruled it is enough for it to contain a single plane. If A is a point on the quadric, any plane thereon through A has to lie in the null prime of A. When $\mathscr{L}$ has the form 2.1 and $\mathbf{B}$ corresponds thereto the null prime of $(1, 1, 1, 1, 1, 1)$ is

$$x_0 + x_1 + x_2 + x_3 + x_4 + x_5 = 0,$$

and the conditions for

$$26.1 \qquad \sum_{i=0}^{5} a_i x_i^2 + \sum_{i<j} x_i x_j = 0$$

to contain the plane $x_0 + x_1 = x_2 + x_3 = x_4 + x_5 = 0$ are

$$a_0 + a_1 = a_2 + a_3 = a_4 + a_5 = 1.$$

Thus 3 squares must be absent from 26.1, 3 present. Again: the null prime of X_0 (see § 2) is

$$x_0 = x_0 + x_1 + x_2 + x_3 + x_4 + x_5,$$

and the conditions for 26.1 to contain the plane $x_1 + x_2 = x_3 + x_4 = x_5 = 0$ are

$$a_0 = 0, \qquad a_1 + a_2 = a_3 + a_4 = 1.$$

The mark a_5 can be either 0 or 1; taking $a_5 = 0$ shows 26.1 to be ruled when 2 squares are present, 4 absent. So $20 + 15 = 35$ ruled quadrics are accounted for.

The ruling would, of course, be patent if 26.1 could be thrown into Plücker's form. For example: if all 6 squares are present 26.1 is

$$(x_0 + x_1 + x_4)(x_0 + x_1 + x_5) + (x_2 + x_3 + x_0)(x_2 + x_3 + x_1) + (x_4 + x_5 + x_2)(x_4 + x_5 + x_3) = 0.$$

References

1. H. F. Baker, *Principles of geometry*, vol. 4 (Cambridge, 1940).
2. ——— *A locus with* 25920 *linear self-transformations* (Cambridge, 1946).
3. G. Castelnuovo, *Ricerche di geometria della retta nello spazio a quattro dimensioni*, Atti del reale Istituto Veneto (7), *2* (1891), 855–901.
4. W. L. Edge, *The geometry of the linear fractional group LF*(4, 2), Proc. Lond. Math. Soc. (3), *4* (1954) 317–42.
5. ——— *The conjugate classes of the cubic surface group in an orthogonal representation*, Proc. Roy. Soc. London, A *233* (1955), 126–46.
6. J. S. Frame, *The classes and representations of the groups of* 27 *lines and* 28 *bitangents*, Ann. Mat. (4), *32* (1951), 83–119.
7. C. M. Hamill, *A collineation group of order* $2^{13}.3^4.5^2.7$, Proc. Lond. Math. Soc. (3), *3* (1953), 54–79.
8. E. Study, *Gruppen zweiseitiger Kollineationen*, Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen (Math.-phys. Klasse) (1912), 453–79.
9. B. L. van der Waerden, *Gruppen von linearen Transformationen* (Berlin, 1935).

University of Edinburgh

POLYTOPES OVER GF(2) AND THEIR RELEVANCE FOR THE CUBIC SURFACE GROUP

H. S. M. COXETER

1. Introduction. In the preceding paper, Edge represented the celebrated "cubic surface group" of order $72.6! = 51840$ as the group of automorphisms of a senary quadratic form over the field of residue-classes mod 2. The object of this sequel is to compare Edge's finite space with a real space, thus identifying his non-ruled quadric in PG(5, 2) with a modular counterpart of the semi-regular polytope 2_{21} which was discovered by Gosset in 1897.

2. The automorphisms of a real quadratic form. The positive definite senary quadratic form

$$E_6 = (x^1)^2 - x^1x^2 + (x^2)^2 - x^2x^3 + (x^3)^2 - x^3x^4 + (x^4)^2 - x^4x^5 + (x^5)^2 - x^3x^6 + (x^6)^2$$

is conveniently symbolized by the "graph"

in which the nodes represent the "square" terms while the branches represent the "product" terms **(4**, pp. 192, 297**)**. The form evidently possesses six involutory automorphisms (or "automorphs") $R_1, \ldots, R_6$, defined as follows. Suppose the kth node of the graph is joined to the ith, jth, etc. Then R_k leaves all the x's unchanged except x^k, which it transforms into

$$-x^k + x^i + x^j + \ldots;$$

for instance, R_3 is

$$(x^1, x^2, x^3, x^4, x^5, x^6) \rightarrow (x^1, x^2, x^2 - x^3 + x^4 + x^6, x^4, x^5, x^6).$$

The six R's generate a group of order 51840 which is the whole group of automorphisms (omitting the trivial automorphism that replaces every x^i by $-x^i$) **(5**, p. 422**)**. To identify this with the group of the 27 lines on the general cubic surface, we observe that its generators satisfy the relations $R_k^2 = 1$ and either

$$R_iR_jR_i = R_jR_iR_j \quad \text{or} \quad R_iR_j = R_jR_i$$

according as the ith and jth nodes of the graph are, or are not, joined. These

Received November 26, 1958.

same relations are satisfied by certain permutations of the 27 lines on the cubic surface (3, p. 458), namely:

$$R_1 = (1\ 2) = \begin{pmatrix} a_1 & b_1 & c_{23} & c_{24} & c_{34} & c_{26} \\ a_2 & b_2 & c_{13} & c_{14} & c_{15} & c_{16} \end{pmatrix},$$

$$R_2 = (2\ 3),\ R_3 = (3\ 4),\ R_4 = (4\ 5),\ R_5 = (5\ 6),$$

$$R_6 = \begin{pmatrix} c_{23} & c_{13} & c_{12} & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & c_{56} & c_{46} & c_{45} \end{pmatrix}.$$

The Diophantine equation

$$E_6 = 1$$

has 36 pairs of solutions, which are easily derived from any one of them, say (1, 0, 0, 0, 0, 0) or

$$\begin{matrix} 10000 \\ 0 \end{matrix},$$

by applying the antomorphisms; for example, by applying R_2, R_3, R_4, R_5, R_6 successively we obtain

$$\begin{matrix} 11000 \\ 0 \end{matrix} \quad \begin{matrix} 11100 \\ 0 \end{matrix} \quad \begin{matrix} 11110 \\ 0 \end{matrix} \quad \begin{matrix} 11111 \\ 0 \end{matrix} \quad \begin{matrix} 11111 \\ 1 \end{matrix}.$$

When regarded as points in an affine 6-space with a Euclidean metric given by the form E_6, these 72 solutions are the vertices of Elte's six-dimensional uniform polytope 1_{22} (1, p. 372; 2, p. 472; 3, p. 470; 4, p. 210; 5, p. 414).

3. The transition from the real field to the field with two elements. Most of the above remarks remain valid if we regard the x's as marks of the field GF(2), that is, if we work in the finite affine 6-space EG(6, 2). Since $-1 = 1$, the minus signs in E_6 can now be replaced by plus signs, and the description of the generating automorphisms becomes still simpler: R_k "increases" the co-ordinate x^k by the sum of its (one, two, or three) neighbours. Another simplification is that the 36 pairs of opposite solutions of $E_6 = 1$ become 36 single solutions which we can identify with the 36 vertices of an "elliptic" 1_{22}, or with the 36 double-sixes on the cubic surface.

Since GF(2) has only the two elements 0, 1, every point of EG(6, 2) that does not satisfy $E_6 = 1$ must satisfy $E_6 = 0$. One such point is the origin. The remaining 27 are the vertices of Gosset's six-dimensional polytope 2_{21} (3, p. 466; 4, p. 202) for which it is natural to use the same symbols that Schläfli invented for the 27 lines. In fact, the point

$$\begin{matrix} 01000 \\ 1 \end{matrix}$$

must be a_1, since it is invariant under all the generating automorphisms

except R_1; and by applying these automorphisms in an appropriate order, we obtain the whole set of 27 as follows:

a_1	a_2	a_3	a_4	a_5
01000	11000	10000	10100	10110
1	1	1	1	1

a_6	c_{56}	c_{46}	c_{36}
10111	10100	10110	10010
1	0	0	0

c_{26}	c_{16}	c_{15}	c_{14}	c_{13}
11010	01010	01011	01001	01101
0	0	0	0	0

c_{12}	c_{23}	c_{24}	c_{25}
00101	11101	11001	11011
0	0	0	0

c_{35}	c_{45}	c_{34}	b_1	b_2
10011	10111	10001	11101	01101
0	0	0	1	1

b_3	b_4	b_5	b_6
00101	00001	00011	00010
1	1	1	1

By regarding the six x's as homogeneous co-ordinates, we thus obtain a partition of the $2^6 - 1 = 63$ points of the finite projective 5-space PG(5, 2) into 27 which lie on the quadric $E_6 = 0$ and 36 which do not. In other words:

In PG(5, 2), *a non-ruled quadric is a polytope* 2_{21}, *and the rest of the space is the "semi-reciprocal" polytope* 1_{22}.

4. The polytope 2_{21}. It is known **(1**, p. 414; **3**, p. 465; **4**, pp. 202, 203**)** that the uniform polytope 2_{21} has 27 vertices

$$a_1, \ldots, a_6; b_1, \ldots, b_6; c_{12}, \ldots, c_{56},$$

216 edges such as a_1a_2; and 720 triangular faces such as $a_1a_2a_3$. Referring to Table II of Edge **(6**, p. 643**)**, we readily identify these with his 27 points m, 216 lines c, and 720 planes j.

The polytope has five-dimensional cells of two kinds: 36 pairs of opposite simplexes $2_{20} = \alpha_5$ such as

$$a_1\ a_2\ a_3\ a_4\ a_5\ a_6, \qquad b_1\ b_2\ b_3\ b_4\ b_5\ b_6$$

and 27 cross-polytopes $2_{11} = \beta_5$ such as

$$a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5$$
$$c_{16} \quad c_{26} \quad c_{36} \quad c_{46} \quad c_{56}.$$

Edge **(6**, p. 626**)** has made the important observation that, when any one of the 72 simplexes α_6 is used as simplex of reference, the quadric has the symmetrical equation

$$\sum_{i<j} x_i x_j = 0.$$

When the 6 + 6 vertices of a pair of opposite simplexes have been removed, the remaining 15 vertices belong to a polytope $0_{31} = t_1\alpha_5$ **(1**, pp. 360, 372**)** such as

$$c_{12}\, c_{13} \ldots c_{56},$$

lying in one of Edge's 36 hyperplanes P.

Other central sections of 2_{21} include 45 triangles such as

$$a_1\ b_2\ c_{12}$$

(Edge's 45 lines g), which correspond to the 45 triangles on the cubic surface, and 120 triangular double-prisms $\{3\} \times \{3\}$ **(3**, pp. 463, 466**)** such as

$$\begin{array}{lll} c_{23} & c_{15} & c_{46} \\ a_2 & b_1 & c_{12} \\ b_3 & a_5 & c_{35} \end{array}$$

(Edge's 120 spaces k), which correspond to the 120 pairs of Steiner trihedra. The latter fall interestingly into 40 sets of three **(3**, p. 466**)**.

5. The polytope 1_{22}. The "elliptic" 1_{22}, which differs from the ordinary 1_{22} **(1**, p. 414**)** in that pairs of opposite elements are identified, has 36 vertices. In PG(5, 2), these are Edge's 36 points p. Each is the centre from which two opposite simplexes α_5 of 2_{21} are in perspective; for example, the simplexes $a_1\ a_2\ a_3\ a_4\ a_5\ a_6$ and $b_1\ b_2\ b_3\ b_4\ b_5\ b_6$ of § 4 are evidently in perspective from the vertex

$$\begin{array}{c} 10101 \\ 0 \end{array}$$

of 1_{22}.

The principal central sections of the ordinary 1_{22} are given in Table IV of **(2**, p. 476**)**. These, with their numbers inserted, are as follows:

36	$\{\ \}$	36	$e\alpha_5$
120	$\{6\}$	120	$\{6\} + \{6\}$
270	$\begin{Bmatrix} 3 \\ 4 \end{Bmatrix}$	270	$\{4\}$
45	$\{3, 4, 3\}$		
216	$e\alpha_4$		
27	$t_1\beta_5$		

The "36 $\{\ \}$" are merely the diagonals joining pairs of opposite vertices, and so correspond to Edge's 36 points p. The 120 hexagons (which reduce to

triangles when opposites are identified) are his 120 lines s. The 270 cuboctahedra (which reduce to complete quadrilaterals) are his 270 planes h. The 45 $\{3, 4, 3\}$ (which reduce to desmic triads of tetrahedra) are his 45 spaces χ. The 216 $e\alpha_4$ (which reduce to complete pentahedra) are his 216 spaces λ. The 27 $t_1\beta_5$ **(1,** pp. 354-360**)** are his 27 spaces M. The 36 $e\alpha_5$ **(1,** p. 366**)** are his 36 spaces P. The 120 pairs of hexagons in completely orthogonal planes are his 120 spaces κ. Finally, the 270 squares (which reduce to point pairs) are his 270 lines t.

References

1. H. S. M. Coxeter, *The polytopes with regular-prismatic vertex figures*, Philos. Trans. Roy. Soc. London, Ser. A, *229* (1930), 329–425.
2. ——— *Finite groups generated by reflections and their subgroups generated by reflections*, Proc. Cambridge Philos. Soc., *30* (1934), 466–82.
3. ——— *The polytope 2_{21} whose 27 vertices correspond to the lines on the general cubic surface*, Amer. J. Math., *62* (1940), 457–86.
4. ——— *Regular polytopes* (London, 1948).
5. ——— *Extreme forms*, Can. J. Math., *3* (1951), 391–441.
6. W. L. Edge, *Quadrics over* GF(2) *and their relevance for the cubic surface group*, Can. J. Math *11* (1959), 631–51.

University of Toronto

ON THE STRUCTURE OF HALF-GROUPS

JAMES V. WHITTAKER

1. Introduction. Furstenberg **(1)** and the author **(4)**, among others, have exhibited postulate systems for groups in terms of the operation $x - y = x + (-y)$. Furstenberg also investigated a system obtained by removing one of his postulates which defines what he called a half-group. A structure theorem for half-groups was given in **(1)**. In the present paper, we prove another structure theorem for half-groups. A more restricted entity, called a pseudo-group, is introduced, and its structure, together with that of a half-group satisfying the left and right cancellation laws, is studied. Finally, some topological questions concerning their structure are also considered.

2. Pseudo-groups. Let G be a set of elements over which is defined a binary operation $x - y$ which satisfies, for any $x, y, z \in G$, the following conditions:

(i) $x - y \in G$,

(ii) There is an $e \in G$ such that $x - y = e$ if, and only if, $x = y$,

(iii) $(x - z) - (y - z) = x - y$.

It was shown in **(4)** that, in terms of the operation $x + y = x - (e - y)$, G is a group. It was also shown that if G satisfies (i), (ii), and

(iv) $(x - z) - (x - y) = y - z$,

then G is a commutative group.

We define a half-group to be a set with an operation satisfying (i) and (iii). A half-group which also has the property

(v) There is an $e \in G$ such that $x - x = e$

will be called a pseudo-group. Before proceeding to a discussion of half-groups, we shall consider briefly the structure of pseudo-groups.

We define a set S to be an extension of a group G if $G \subset S$ and if there exists a single-valued function f defined on S such that $f(x) \in G$ for all $x \in S$ and $f(u) = u$ for all $u \in G$. The set S becomes a pseudo-group under the operation

$$x \circ y = f(x) - f(y), \qquad x, y \in S. \tag{1}$$

Indeed, $x \circ y \in G \subset S$, and $x \circ x = f(x) - f(x) = e$. Finally,

$$\begin{aligned}(x \circ z) \circ (y \circ z) &= f(f(x) - f(z)) - f(f(y) - f(z)) \\ &= (f(x) - f(z)) - (f(y) - f(z)) = f(x) - f(y) = x \circ y,\end{aligned}$$

and (i), (iii), and (v) are satisfied.

THEOREM 1. *Every pseudo-group S contains a group G such that S is an extension of G and the operation in S satisfies* (1).

Received October 6, 1958.

Proof. Let $x - y$ be the operation in S, and set $f(x) = x - e$. We will show that S is an extension of the group $G = f(S)$. For any $x, y \in S$, we have

$$(x - y) - e = (x - y) - (y - y) = x - y.$$

Setting $y = e$, we obtain $f(f(x)) = f(x)$, whence $f(u) = u$ for all $u \in G$. It also follows that if $u, v \in G$, then $u - v = f(u - v) \in G$, and G satisfies (i). Since (iii) and (v) hold in S, they hold also in G. To verify the remaining half of (ii), we observe that if $u, v \in G$ and $u - v = e$, then, since $f(u) = u$ and $f(v) = v$,

$$\begin{aligned} u = u - e &= (u - v) - (e - v) = e - (e - v) \\ &= (v - v) - (e - v) = v - e = v. \end{aligned}$$

Hence, G is a group. Finally, it follows from (iii) that, for all $x, y \in S$, $x - y = f(x) - f(y)$, and the proof is complete.

In any pseudo-group S, we may introduce the sum $x + y = x - (e - y)$ for all $x, y \in S$. Then

$$x + y = f(x) - f(e - y) = f(x) - (e - f(y)) = f(x) + f(y).$$

Hence, this operation is associative. A pseudo-group is said to be commutative if $x + y = y + x$.

THEOREM 2. *If S is a set with an operation satisfying* (i) *and* (iv), *then S is a commutative pseudo-group.*

Proof. For any $x, y, z \in S$, we have

$$\begin{aligned} (x - z) - (y - z) &= ((z - z) - (z - x)) - ((z - z) - (z - y)) \\ &= (z - y) - (z - x) = x - y, \end{aligned}$$

which proves (iii). Next, $x - x = (x - y) - (x - y) = y - y = e$. Since the group $G = S - e$ also satisfies (iv), G is commutative. From the remark preceding Theorem 2, it follows that $x + y = y + x$ for all $x, y \in S$.

In terms of subtraction, the left and right cancellation laws take the following forms: for all $x, y, z \in S$,

(vi) $x - y = x - z$ implies $y = z$,

(vii) $x - z = y - z$ implies $x = y$.

The next theorem shows, incidentally, that (vi) implies (vii) in a half-group. That (vii) does not imply (vi) can be seen by considering what we shall call the simple half-group, a set in which $x - y$ is defined to be x. The structure of half-groups satisfying (vii) will be taken up in the next section.

THEOREM 3. *A half-group in which the left cancellation law holds is a group.*

Proof. Suppose that S satisfies (i), (iii), and (vi). For any $u, x, y, z \in S$, we have

$$\begin{aligned} (u - x) - z &= ((u - x) - (x - x)) - (z - (x - x)) \\ &= (u - x) - (z - (x - x)). \end{aligned}$$

Hence, $z - (x - x) = z = z - (y - y)$, and $x - x = y - y = e$. If $x - y = e = x - x$, then $x = y$ and (ii) is satisfied. Therefore, G is a group.

3. Half-groups. Following Furstenberg **(1)**, we introduce the following definitions. We call an element a of a half-group H an idempotent if $a - a = a$. A subset $K \subset H$ is called invariant in H if $x - (x - K) \subset K$ for all $x \in H$. If K is a sub-half-group of H which contains all idempotents in H, then the relation $x \equiv y$ defined by $x - y \in K$ is an equivalence relation. In fact, $x \equiv x$ since $x - x$ is an idempotent. If $x \equiv z$ and $y \equiv z$, then $x - y = (x - z) - (y - z) \in K$, and $x \equiv y$. The resulting quotient space will be denoted by H/K. A weaker form of the following theorem was proved in **(1**, Theorem 4**)**.

THEOREM 4. *In the half-group H, let K be an invariant sub-half-group containing all idempotents of H. Then H/K is a group.*

Proof. We refer to **(1)** for the proof that if $u, v, x, y \in H$ and $u \equiv x$, $v \equiv y$, then $u - v \equiv x - y$. Thus, if x^* and y^* are the equivalence classes containing x and y, respectively, then $(x - y)^* = x^* - y^*$. Clearly H/K satisfies (i) and (iii). If e^* is the equivalence class containing K, then $x^* - x^* = e^*$, since $x - x$ is an idempotent. Finally, if $x^* - y^* = e^*$, then $x^* = y^*$, and (ii) is satisfied. Hence, H/K is a group.

Let G be a group and $\{S_\alpha\}_{\alpha \in A}$ a family of disjoint sets, indexed by the set A, with the following properties:

(a) For each $\alpha \in A$, there exists a single-valued function f_α mapping S_α onto G,

(b) For each $\alpha, \beta \in A$, there exists a one-to-one function $\phi_{\alpha\beta}$ mapping G into S_α such that $f_\alpha(\phi_{\alpha\beta}(g)) = g$ for each $g \in G$.

We now define an operation $x \circ y$ in $H = \bigcup S_\alpha$ as follows: if $x \in S_\alpha$ and $y \in S_\beta$, then

$$x \circ y = \phi_{\alpha\beta}(f_\alpha(x) - f_\beta(y)). \tag{2}$$

We note that $x \circ y \in S_\alpha$. Under this operation, H becomes a half-group. Since (i) is clearly satisfied, we have only to verify (iii). If $z \in S_\gamma$, we have

$$\begin{aligned}(x \circ z) \circ (y \circ z) &= \phi_{\alpha\beta}\{f_\alpha(\phi_{\alpha\gamma}(f_\alpha(x) - f_\gamma(z))) - f_\beta(\phi_{\beta\gamma}(f_\beta(y) - f_\gamma(z)))\} \\ &= \phi_{\alpha\beta}((f_\alpha(x) - f_\gamma(z)) - (f_\beta(y) - f_\gamma(z))) \\ &= \phi_{\alpha\beta}(f_\alpha(x) - f_\beta(y)) = x \circ y.\end{aligned}$$

Each S_α becomes a pseudo-group in which $\phi_{\alpha\alpha}f_\alpha$ plays the role of f in the representation of pseudo-groups given in § 2. If e is the identity of G, then $\phi_{\alpha\alpha}(e) = e_\alpha$ is the identity in S_α, since $x \circ x = e_\alpha$ for all $x \in S_\alpha$. The idempotents in H are precisely the elements e_α for all $\alpha \in A$. It follows from (2) that $\phi_{\alpha\beta}(f_\alpha(x)) = x \circ e_\beta$ for all $x \in S_\alpha$. In particular, $S_\alpha \circ e_\alpha$ is isomorphic to G.

THEOREM 5. *Every half-group H can be represented as the union of disjoint pseudo-groups in which the operation is defined by* (2).

Proof. Let $E = \{e_\alpha\}_{\alpha \in A}$ be the set of idempotents of H indexed by the set A, and, for each $\alpha \in A$, let $S_\alpha = \{x \in H: x - x = e_\alpha\}$. If $x \in S_\alpha$ and $y \in H$, then clearly $x - y \in S_\alpha$, and S_α satisfies (i). Since (iii) holds in H, it holds also in S_α. Finally, S_α satisfies (v) with $e = e_\alpha$, so that S_α is a pseudo-group. In addition, $\alpha \neq \beta$ implies $S_\alpha \cap S_\beta = \emptyset$, and $H = \cup S_\alpha$.

We turn now to the construction of G and the functions appearing in (2). For any $x, y \in H$ and $a \in E$, we have

$$\begin{aligned}(x - a) - y &= ((x - a) - (a - a)) - (y - (a - a)) \\ &= (x - a) - (y - a) = x - y.\end{aligned}$$

Similarly, if $b \in E$, then $x - (y - b) = x - y$. Combining these relations, we obtain

$$(x - a) - (y - b) = x - y. \tag{3}$$

(For the sake of compactness, we shall also let (3) stand for the two preceding relations.) If we set $F = E - E = \{u - v: u, v \in E\}$, then F is a sub-half-group by virtue of (3). Repeated application of (3) also shows that $x - (x - z) = x - x \in F$ for all $x \in H$ and $z \in F$. Hence, by Theorem 4, $H/F = G$ is a group. The corresponding equivalence relation will be written $x \equiv y$, and the quotient mapping from H onto G will be denoted by h.

For each $\alpha \in A$, let $G_\alpha = S_\alpha - e_\alpha$. We now define functions $\psi_\alpha: H \to G_\alpha$ as follows:

$$\psi_\alpha(x) = e_\alpha - (e_\alpha - x).$$

For any $x \in H$, we have $e_\alpha - x \in S_\alpha$. From Theorem 1 it follows that the difference of any two elements in S_α lies in $S_\alpha - e_\alpha = G_\alpha$, whence $\psi_\alpha(x) \in G_\alpha$. In addition,

$$h(\psi_\alpha(x)) = h(e_\alpha) - (h(e_\alpha) - h(x)) = e - (e - h(x)) = h(x),$$

where e is the identity of G. Hence, $x \equiv \psi_\alpha(x)$ and

$$x - \psi_\alpha(x) = e_\gamma - e_\delta. \tag{4}$$

If $x \in S_\beta$, then $x - \psi_\alpha(x) \in S_\beta$ and $e_\gamma - e_\delta \in S_\gamma$, whence $\gamma = \beta$. (Further calculation shows that $\delta = \alpha$, but we shall not need this fact.) Thus every equivalence class intersects each G_α. Furthermore, no two elements of G_α are equivalent, that is, $G_\alpha \cap F = \{e_\alpha\}$. In fact, if $y \in G_\alpha$ and $y = e_\alpha - e_\beta$, then $y = y - e_\alpha = (e_\alpha - e_\beta) - e_\alpha = e_\alpha$. Hence, the restriction of h to G_α is an isomorphism h_α between G_α and G.

The functions f_α are now defined to be the restrictions of h to S_α, and the $\phi_{\alpha\beta}$ are defined, for each $\alpha, \beta \in A$, and each $g \in G$, by

$$\phi_{\alpha\beta}(g) = h_\alpha^{-1}(g) - e_\beta.$$

Since $x - e_\beta \equiv x$ for all $x \in S_\alpha$ because of (3), it follows that $f_\alpha(\phi_{\alpha\beta}(g)) = g$ for all $g \in G$, and conditions (a) and (b) are fulfilled. Moreover, if $x \in S_\alpha$, then $h_\alpha^{-1}(f_\alpha(x)) = x - e_\alpha$, for each member is equivalent to x and lies in

G_α, and such elements in G_α must be identical. Similarly, if $y \in S_\beta$, then $h_\alpha^{-1}(f_\beta(y)) = \psi_\alpha(y)$, for each member lies in G_α and is equivalent to y. Finally, if $x \in S_\alpha$ and $y \in S_\beta$, then, using (3) and (4),

$$\begin{aligned} x - y &= (x - \psi_\alpha(y)) - (y - \psi_\alpha(y)) \\ &= ((x - e_\alpha) - \psi_\alpha(y)) - (e_\beta - e_\delta) \\ &= (h_\alpha^{-1}(f_\alpha(x)) - h_\alpha^{-1}(f_\beta(y))) - e_\beta \\ &= h_\alpha^{-1}(f_\alpha(x) - f_\beta(y)) - e_\beta = \phi_{\alpha\beta}(f_\alpha(x) - f_\beta(y)). \end{aligned}$$

This completes the proof.

COROLLARY. *Let $K \supset E$ be an invariant sub-half-group of H. Then there exists an invariant subgroup N of G such that $K/F \cong N$ and $H/K \cong G/N$.*

Proof. We adopt the notation of the proof of Theorem 5. Let $h(K) = N$. Then N is clearly an invariant subgroup of G. Since the restriction of h to K is the canonical mapping from K onto K/F, N and K/F are isomorphic. Now let $L = h^{-1}(N)$ and k be the canonical mapping from H onto H/K. Clearly $x - y \in K$ implies $x - y \in L$ for all $x, y \in H$. Conversely, if $x - y \in L$, then there exists a $z \in K$ such that $x - y \equiv z \pmod F$. Since $K \supset F$, $x - y \equiv z \pmod K$ and

$$k(x) - k(y) = k(x - y) = k(z) = \bar{e},$$

where $\bar{e}$ is the identity in H/K. Hence, $k(x) = k(y)$ and $x - y \in K$.

Let θ be the canonical mapping from G onto G/N. For any $x, y \in H$, the relation $\theta(h(x)) = \theta(h(y))$ is evidently equivalent to $h(x) - h(y) \in N$ and, hence, to $x - y \in L$. Thus, the correspondence between H/K and G/N defined by $x^* \leftrightarrow \theta(h(x))$, where x^* is the equivalence class (mod K) containing x, is one-to-one and onto. Finally,

$$x^* - y^* = (x - y)^* \leftrightarrow \theta(h(x - y)) = \theta(h(x)) - \theta(h(y)),$$

and the correspondence is a group isomorphism.

Let $\{G_\alpha\}_{\alpha \in A}$ be a family of disjoint groups, indexed by the set A, satisfying the following condition:

(c) For each $\alpha, \beta \in A$, there exists a unique isomorphism $\theta_{\alpha\beta}: G_\beta \to G_\alpha$. In addition, for all $\alpha, \beta, \gamma \in A$, $\theta_{\alpha\beta} \cdot \theta_{\beta\gamma} = \theta_{\alpha\gamma}$.

It follows from (c) that $\theta_{\alpha\alpha}(x) = x$ for all $x \in G_\alpha$. We now define an operation $x \circ y$ in $H = \cup G_\alpha$ as follows: if $x \in G_\alpha$ and $y \in G_\beta$, then

$$x \circ y = x - \theta_{\alpha\beta}(y). \tag{5}$$

Thus $x \circ y \in G_\alpha$. If $z \in G_\gamma$, then

$$\begin{aligned} (x \circ z) \circ (y \circ z) &= (x - \theta_{\alpha\gamma}(z)) - \theta_{\alpha\beta}(y - \theta_{\beta\gamma}(z)) \\ &= (x - \theta_{\alpha\gamma}(z)) - (\theta_{\alpha\beta}(y) - \theta_{\alpha\gamma}(z)) = x \circ y. \end{aligned}$$

Hence, H is a half-group. Furthermore, if $x \circ z = y \circ z$, then $\alpha = \beta$ and $x = y$, whence H satisfies (vii).

We can define a second operation $x \oplus y$ in H as follows: $x \oplus y = x \circ (e_\gamma \circ y)$, where e_γ is the identity in G_γ. We have

$$(6) \quad x \oplus y = x - \theta_{\alpha\gamma}(e_\gamma - \theta_{\gamma\beta}(y)) = x - (e_\alpha - \theta_{\alpha\beta}(y)) = x + \theta_{\alpha\beta}(y).$$

Thus $x \oplus y$ is independent of γ, and it is easily verified that this operation is associative. In addition, each e_γ is a right identity of H.

THEOREM 6. *If H is a half-group in which the right cancellation law holds, then H can be represented as the union of disjoint, isomorphic groups in which the operation is defined by* (5). *Moreover, H is a semigroup under the operation defined by* (6).

Proof. We adopt the notation of the proof of Theorem 5. It follows from (3) that, for all $x \in H$ and $\alpha \in A$, $x - e_\alpha = (x - e_\alpha) - e_\alpha$, whence, by (vii), $x = x - e_\alpha$. Thus $S_\alpha = G_\alpha$, and f_α is one-to-one. From (b) we infer that $\phi_{\alpha\beta} = f_\alpha^{-1}$ for all $\beta \in A$. If $x \in G_\alpha$ and $y \in G_\beta$, then

$$x - y = f_\alpha^{-1}(f_\alpha(x) - f_\beta(y)) = x - f_\alpha^{-1}(f_\beta(y)).$$

If we set $f_\alpha^{-1} f_\beta = \theta_{\alpha\beta}$, then the conditions of (c) are fulfilled and (5) is verified. The rest of the theorem follows from the discussion preceding it.

COROLLARY. *If H is a half-group satisfying* (vii) *in which the operation $x \oplus y$ is commutative, then H is a (commutative) group.*

Proof. If $x \in G_\alpha$ and $y \in G_\beta$, then $x \oplus y \in G_\alpha$ and $y \oplus x \in G_\beta$, whence $\alpha = \beta$ and $x \oplus y = x + y$.

4. Topological pseudo-groups. A topological pseudo-group S is a pseudo-group with a topology J in which $x - y$ is continuous jointly in x and y. Thus $G = S - e$ is a topological group in its relative topology. This definition, however, places very little restriction on the topology in the complement of G, and it will sometimes be necessary to impose the additional condition that the topology does not separate x and $x - e$:

(viii) For every $U \in J$ and every $x \in S$, $x \in U$ if, and only if, $x - e \in U$.

Let S be an extension of a topological group G with topology $J(G)$, and let f be defined as in § 2. We introduce a topology $J(S)$ into S as follows:

$$(7) \qquad J(S) = \{f^{-1}(U) : U \in J(G)\}.$$

With this topology and the operation defined in (1), S becomes a topological pseudo-group. To show that subtraction is continuous, let $x, y \in S$ and $x \circ y \in W$, where $W \in J(S)$. Then there exist sets $U_0, V_0 \in J(G)$ such that $f(x) \in U_0$, $f(y) \in V_0$, and $U_0 - V_0 \subset f(W)$. Hence, $x \in U = f^{-1}(U_0)$, $y \in V = f^{-1}(V_0)$, and $U \circ V \subset W$. Since $f(x) = x \circ e$, (viii) is also satisfied.

THEOREM 7. *Every topological pseudo-group S which satisfies* (viii) *can be represented as an extension of a topological group in which the operation is defined by* (1) *and the topology by* (7).

Proof. Let $G = S - e$ and $f(x) = x - e$ for $x \in S$. Then G is a topological group in its relative topology, and f is continuous on S. From (viii) it follows that if $U \in J(S)$, then $f^{-1}(U \cap G) = U$, which verifies (7). The rest follows from Theorem 1.

In each of the following corollaries, S denotes a topological pseudo-group satisfying (viii). We define S to be T_i $(i = 1, 2)$ if $S - e$ is T_i in its relative topology. A topological space is T_1 if each of its points is closed, and T_2 if every two distinct points have disjoint neighbourhoods. A space is first countable if every point has a countable base of neighbourhoods.

COROLLARY 1. *If $C \subset S$ is closed, and $x \in S$ is such that $x - e \notin C - e$, then there exists a continuous ϕ: $S \to [0, 1]$ such that $\phi(x) = 0$ and $\phi(y) = 1$ for $y \in C$.*

Proof. Since $f(x) \notin f(C)$ and $f(C)$ is closed in G, there exists (3, p. 188) a continuous ϕ_0: $G \to [0, 1]$ such that $\phi_0(f(x)) = 0$ and $\phi_0 = (f(y)) = 1$ for $y \in C$. Hence, $\phi = \phi_0 f$ has the desired properties.

COROLLARY 2. *If S is first countable, then S is pseudo-metrizable with a right-invariant pseudo-metric.*

Proof. Since G is first countable in its relative topology, it follows from (3, p. 210) that G is pseudo-metrizable with a right-invariant pseudo-metric $r(x, y)$. Then $s(x, y) = r(f(x), f(y))$ is the desired pseudo-metrix for S. Since $f(x - z) = f(x) - f(z)$, s is also right-invariant.

COROLLARY 3. *If S is T_1 and locally compact, then there exists a right-invariant Haar measure μ in S.*

Proof. Since G is T_1 and locally compact, there exists a right-invariant Haar measure μ_0 on the Borel sets of G (cf. 2, p. 254). Each Borel set $B \subset S$ is evidently of the form $f^{-1}(B_0)$ where B_0 is a Borel set relative to G. Hence, $\mu(B) = \mu_0(f(B))$, and the right-invariance follows as in the proof of Corollary 2.

To show the necessity for some additional assumption such as (viii) in the preceding corollaries, we give the following example, based on one in (1), in which S is a topological pseudo-group which does not satisfy (viii). In addition, S is compact, connected, first countable, and T_1, but not T_2, and hence, not pseudo-metrizable. Let S be the interval $[0, 2]$, and let the operation in S be defined as subtraction (mod 1). Then S is evidently a commutative pseudo-group, and $G = [0, 1)$. The neighbourhoods of the points in $(0, 2]$ are those of the relative Euclidean topology, while the neighbourhoods of 0 are of the form $[0, \rho) \cup (\sigma, 1)$, where $\rho, \sigma \in (0, 1)$. Evidently S is a topological pseudo-group, and the other properties of S are easily verified. In particular, 0 and 1 have no disjoint neighbourhoods.

5. Topological half-groups. A topological half-group H is defined as a half-group with a topology in which $x - y$ is jointly continuous in x and y.

For the theorems in this section, we do not need as strong a condition as (viii), and we shall require only the following:

(ix) For every $U \in J(H)$ and every $x \in H$, $x \in U$ implies $x - (x - x) \in U$.

The notation of the proof of Theorem 5 will be adhered to throughout this section. From the foregoing definition, we see that each G_α is a (not necessarily T_1) topological group in its relative topology, and each S_α is a topological pseudo-group.

THEOREM 8. *Let H be a topological half-group satisfying* (ix) *and $K \supset E$ an invariant sub-half-group of H. Then the quotient mapping from H onto H/K becomes an open, continuous mapping when H/K is given the quotient topology, and H/K becomes a topological group.*

Proof. We first consider the case $K = F$ and show that, for any open $U \subset H$, $U^* = h^{-1}(h(U))$ is open. Let $V = \bigcup_{\alpha \in A} \psi_\alpha^{-1}(U)$. Since ψ_α is continuous on H and $\psi_\alpha(x) \equiv x$ for all $x \in H$, V is open and $V \subset U^*$. If $x \in U^*$, then $x \equiv u$ for some $u \in U$. Now $u \in S_\alpha$ for some $\alpha \in A$, and $u \equiv u - e_\alpha = u - (u - u) \in G_\alpha \cap U$ in view of (ix). Thus, we can assume that $u \in G_\alpha$. Then $\psi_\alpha(x) \equiv \psi_\alpha(u) = u$, and since $\psi_\alpha(x) \in G_\alpha$, we have $\psi_\alpha(x) = u$. Hence, $x \in V$ and $U^* \subset V$, so that $U^* = V$ is open. It follows immediately that h is continuous and open, and G is a topological group in the quotient topology.

Now let $K \supset E$ be an arbitrary invariant sub-half-group of H, k the quotient mapping from H onto H/K, and θ the quotient mapping from G onto G/N, where $N = h(K)$. From the corollary to Theorem 5, it follows that, for each $x \in H$, $x^* = k^{-1}(k(x)) = h^{-1}(\theta^{-1}(\theta(h(x))))$. Since h and θ are each open, continuous mappings, we conclude that if $U \subset H$ is open, then $U^* = k^{-1}(k(U))$ is open. Hence, k is open and continuous, and H/K is a topological group in its quotient topology.

A topological half-group H will be said to be T_i $(i = 1, 2)$, provided $\Gamma = \bigcup G_\alpha$ is T_i in its relative topology. If H satisfies (ix), then each h_α is a topological, as well as an algebraic, isomorphism between G_α in its relative topology and G. The continuity of h_α follows directly from Theorem 8. Let $U \subset H$ be open. Since the range of ψ_α is G_α and $\psi_\alpha(x) \equiv x$ for all $x \in H$, each $x \in \psi_\alpha^{-1}(U)$ is congruent (mod F) to some element in $U \cap G_\alpha$. Moreover, $\psi_\alpha(x) = x$ for each $x \in G_\alpha$, so that $\psi_\alpha^{-1}(U) \supset U \cap G_\alpha$. Hence, $h(U \cap G_\alpha) = h(\psi_\alpha^{-1}(U))$ is open, and h_α is open. Thus, if H is T_i, then G is also T_i.

Consider now the set $G \times E$ and the operation defined in it as follows:

$$(g_1, e_\alpha) \circ (g_2, e_\beta) = (g_1 - g_2, e_\alpha).$$

Under this operation, $G \times E$ evidently becomes a half-group satisfying (vii) (cf. the discussion preceding Theorem 6). If we introduce the product topology in $G \times E$ derived from the quotient topology in G and the relative topology from H in E, then $G \times E$ becomes a topological half-group. In fact, if $U_1 \times V_1$ and $U_2 \times V_2$ are open in $G \times E$, then $(U_1 \times V_1) \circ (U_2 \times V_2) = (U_1 - U_2, V_1$.

Let $\phi: H \to G \times E$ be defined as follows: if $x \in S_\alpha$, then $\phi(x) = (h(x), e_\alpha)$. Since, for any $y \in H$, $x - y \in S_\alpha$, we have

$$\phi(x - y) = (h(x) - h(y), e_\alpha) = \phi(x) \circ \phi(y).$$

The restriction of ϕ to Γ is clearly one-to-one and onto, but, in general, $x - y \neq \phi^{-1}(\phi(x) \circ \phi(y))$ for $x, y \in \Gamma$.

THEOREM 9. *If H is a topological half-group satisfying* (ix), *then ϕ is a continuous, open homomorphism of H onto $G \times E$.*

Proof. It remains to prove that ϕ is continuous and open. Let $\omega(x) = x - x$ for all $x \in H$. Then ω is continuous on H, $\omega(S_\alpha) = e_\alpha$ for each $\alpha \in A$, and $\phi(x) = (h(x), \omega(x))$. If $U \times V \subset G \times E$ is open, then $\phi^{-1}(U \times V) = h^{-1}(U) \cap \omega^{-1}(V)$ is open in H, and ϕ is continuous. Now let $U \subset H$ be open and $y \in \phi(U)$. By an argument used in the proof of Theorem 8, we can assume that $y = \phi(x)$, where $x \in U \cap G_\alpha$ for some $\alpha \in A$. Since $x = e_\alpha - (e_\alpha - x)$, there exist neighbourhoods V, W, of x, e_α, respectively, such that $W - (W - V) \subset U$. Let $V' = h(V)$ and $W' = W \cap E$. Then $V' \times W'$ is open in $G \times E$, and clearly $y \in V' \times W'$. If $v \in V$ and $e_\beta \in W$, then $e_\beta - (e_\beta - v) \in U$, $h(e_\beta - (e_\beta - v)) = h(v)$, and $\omega(e_\beta - (e_\beta - v)) = e_\beta$. Hence,

$$(h(v), e_\beta) = \phi(e_\beta - (e_\beta - v)) \in \phi(U),$$

and $V' \times W' \subset \phi(U)$. Therefore, ϕ is open.

A further consequence of the continuity of ω is that, in a T_1 topological half-group H, each S_α is closed in H.

THEOREM 10. *If H is a connected, T_1 topological half-group with a finite number of idempotents, then H is a (topological) pseudo-group.*

Proof. We have $\omega(H) = E$, whence E is connected and T_1 in its relative topology. Since a finite T_1 space is discrete, E reduces to a single point.

COROLLARY. *If, in addition, the right cancellation law holds in H, then H is a (topological) group.*

To show the necessity for the assumption that E be finite, let $H = [0, 1]$ with the relative Euclidean topology and the structure of the simple half-group mentioned in § 2. Evidently H is a compact, connected, T_1 topological half-group in which every element is an idempotent.

REFERENCES

1. H. Furstenberg, *The inverse operation in groups*, Proc. Amer. Math. Soc., *6* (1955), 991–7.
2. P. R. Halmos, *Measure theory* (New York, 1950).
3. J. L. Kelley, *General topology* (New York, 1955).
4. J. V. Whittaker, *On the postulates defining a group*, Amer. Math. Monthly, *62* (1955), 636–40.

University of British Columbia

ON THE CLASS NUMBER OF REPRESENTATIONS OF AN ORDER

IRVING REINER

1. Introduction. We shall use the following notation throughout:

R = Dedekind ring **(5)**.
$\mathfrak{u}$ = multiplicative group of units in R.
h = class number of R.
K = quotient field of R.
$\mathfrak{p}$ = prime ideal in R.
$R_{\mathfrak{p}}$ = ring of $\mathfrak{p}$-adic integers in K.

We assume that h is finite, and that for each prime ideal $\mathfrak{p}$, the index $(R:\mathfrak{p})$ is finite.

Let A be a finite-dimensional separable algebra over K, with an identity element e **(4**, p. 115**)**. Let G be an R-order in A, that is, G is a subring of A satisfying

(i) $e \in G$,
(ii) G contains a K-basis of A,
(iii) G is a finitely-generated R-module.

By a G-module we shall mean a left G-module which is a finitely-generated torsion-free R-module, on which e acts as identity operator. An A-module is defined analogously, replacing R by K. We shall assume, unless otherwise stated, that K is a splitting field for A; thus, the only possible A-endomorphisms of an irreducible A-module X are the scalar multiplications $x \to \alpha x$, $x \in X$, where $\alpha \in K$.

As in **(3)**, we may form the non-zero ideal $\mathfrak{g} \subset R$, defined as the intersection of the ideals which annihilate the one-dimensional cohomology groups $H(G, T)$, where T ranges over the set of two-sided G-modules. (In the special case where $G = R(\Pi)$ is the group ring of a finite group Π, the ideal $\mathfrak{g}$ is the principal ideal generated by the group order $(\Pi : 1)$.) Let $P = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_t\}$ be the set of distinct prime divisors of $\mathfrak{g}$, and set

$$\mathfrak{g} = \prod_{\mathfrak{p} \in P} \mathfrak{p}^{\gamma(\mathfrak{p})}. \tag{1}$$

For any G-module M, let KM be the A-module which consists of the K-linear combinations of the elements of M. If we set $A_{\mathfrak{p}} = R_{\mathfrak{p}}G$, we may likewise define the $A_{\mathfrak{p}}$-module $M_{\mathfrak{p}} = R_{\mathfrak{p}}M$. Two G-modules M and N are said to be in the same *genus* (notation: $M \vee N$) if and only if for each $\mathfrak{p}$, the modules

Received June 12, 1958. This research was supported in part by the Office of Naval Research.

M_p and N_p are A_p-isomorphic. As is shown in **(7)**, $M \vee N$ if and only if $KM \cong KN$ and $M_p \cong N_p$ for each $p \in P$.

For any A-module L', let $S(L')$ be the collection of G-modules L for which $KL \cong L'$. Suppose that $S(L')$ splits into r_g genera, and into r_G classes under G-isomorphism. As is shown in **(6; 7;** and **9)**, both r_g and r_G are finite. The purpose of this paper is to consider the relation between r_g and r_G. For the special case where L' is irreducible, Maranda **(7)** has shown that $r_G = hr_g$. We shall restrict ourselves to the case where the irreducible constituents of L' are distinct from one another. If L' has k distinct irreducible constituents, we shall prove

$$r_G \geqslant h^k r_g. \tag{2}$$

Further, we shall show that equality holds provided that

(3) For each $\alpha \in R$ such that $(\alpha) + \mathfrak{g} = R$, there exists $\beta \in \mathfrak{u}$ for which $\beta \equiv \alpha \pmod{\mathfrak{g}^{k-1}}$.

Finally, we shall obtain formulas for r_g and r_G in the special case where $k = 2$. These formulas will show that if condition (3) fails, then r_G may exceed $h^2 r_g$ for this case.

2. Binding homomorphisms. In this section, we shall drop the hypothesis that K is a splitting field for the algebra A. Let L be a G-module which contains a submodule M, and assume that M is an R-direct summand of L. Define $N = L/M$ to be the factor G-module. Every element of L is then uniquely representable as an ordered pair (n, m), $n \in N$, $m \in M$, where the structure of L as R-module is given by

$$(n, m) + (n', m') = (n + n', m + m'), \qquad \alpha(n, m) = (\alpha n, \alpha m), \tag{4}$$

for $n, n' \in N$, $m, m' \in M$, $\alpha \in R$. Further, the action of G on L is given by

$$g(n, m) = (gn, \Lambda_g(n) + gm), \qquad g \in G, \text{ where } \Lambda_g \in \mathrm{Hom}_R(N, M). \tag{5}$$

Let $\Lambda : G \to \mathrm{Hom}_R(N, M)$ be the R-homomorphism defined by $g \to \Lambda_g$. The condition $(gh)(n, m) = g(h(n, m))$ is equivalent to

$$\Lambda_{gh}(n) = g\Lambda_h(n) + \Lambda_g(hn), \qquad g, h \in G, n \in N. \tag{6}$$

Call $\Lambda \in \mathrm{Hom}_R(G, \mathrm{Hom}_R(N, M))$ a *binding homomorphism* if **(6)** holds, and let $B(N, M)$ be the R-module consisting of all binding homomorphisms relative to N, M. The R-G-module L is then completely determined by equations (4) and (5), once an element $\Lambda \in B(N, M)$ is fixed. Let us denote this module L by $(N, M; \Lambda)$.

It is convenient to turn $\mathrm{Hom}_R(N, M)$ into a two-sided G-module T by defining

$$(gt)(n) = g(t(n)), \quad (tg)n = t(gn), \quad g \in G, n \in N, t \in \mathrm{Hom}_R(N, M).$$

We may then characterize $B(N, M)$ as the set of all $\Lambda \in \mathrm{Hom}_R(G, T)$ for which

$$\Lambda_{gh} = g\Lambda_h + \Lambda_g h, \qquad g, h \in G. \tag{7}$$

Now fix $t \in T$, and define $\Lambda \in \mathrm{Hom}_R(G, T)$ by

$$\Lambda_g = gt - tg, \qquad g \in G.$$

We find readily that $\Lambda \in B(N, M)$. Let $B'(N, M)$ be the R-module consisting of all the binding homomorphisms so obtained by letting t range over all elements of T. Define the R-module

$$C(N, M) = B(N, M)/B'(N, M).$$

From **(9)** we know that $C(N, M)$ contains only finitely many elements. Furthermore, from the definition of the ideal $\mathfrak{g}$, we have

$$\mathfrak{g} \cdot B(N, M) \subset B'(N, M)$$

for any N, M. Finally, if $[\Lambda]$ denotes the class $\Lambda + B'(N, M)$ of the element $\Lambda \in B(N, M)$, then we have:

$$[\Lambda] = [\Lambda'] \Rightarrow (N, M; \Lambda) \cong (N, M; \Lambda').$$

In fact, if $t \in T$ is such that $\Lambda_g' - \Lambda_g = gt - tg$, $g \in G$, then the map $(n, m) \to (n, m - tn)$ gives the desired isomorphism.

In the above discussion, replace R by R_p. If L^* is an A_p-module which contains a submodule M^* as R_p-direct summand, then $L^* = (N^*, M^*; \Lambda^*)$, where $N^* = L^*/M^*$, and where

$$\Lambda^* : A_p \to \mathrm{Hom}_{R_p}(N^*, M^*).$$

is an R_p-homomorphism satisfying $\Lambda^*_{xy} = x\Lambda^* y + \Lambda^*_x y$, $x, y \in A_p$. Define $B(N^*, M^*)$, $B'(N^*, M^*)$ and $C(N^*, M^*)$ as above. For $\Lambda^* \in B(N^*, M^*)$, again let $[\Lambda^*] = \Lambda^* + B'(N^*, M^*)$. If $\gamma(p)$ is defined as in (1), we have

$$\pi^{\gamma(p)} B(N^*, M^*) \subset B'(N^*, M^*) \tag{8}$$

where π is an element of p such that $\pi \notin p^2$.

Now let N, M be G-modules, and let N_p, M_p be the corresponding A_p-modules. There is a natural isomorphism of $B(N, M)$ into $B(N_p, M_p)$ which may be described as follows: for each $\Lambda \in B(N, M)$ and each $g \in G$, the map $\Lambda_g \in \mathrm{Hom}_R(N, M)$ may be extended in a unique manner to an element of $\mathrm{Hom}_{R_p}(N_p, M_p)$; we may then define Λ_x for each $x \in A_p$ by linearity. In this way, Λ is extended in a unique manner to an element $\Lambda^p \in B(N_p, M_p)$. The map $\Lambda \to \Lambda^p$ carries $B'(N, M)$ into $B'(N_p, M_p)$, and so induces an R-homomorphism of $C(N, M)$ into $C(N_p, M_p)$.

We may now define an R-homomorphism

$$\phi : \quad C(N, M) \to \sum_{p \neq P} C(N_p, M_p)$$

by means of

$$\phi[\Lambda] = ([\Lambda^{p_1}], \ldots, [\Lambda^{p_t}]).$$

From **(8)**, we know that ϕ has kernel 0. We shall in fact show that ϕ is an isomorphism "onto."

THEOREM 1.

$$C(N, M) \cong \sum_{p \in P} C(N_p, M_p).$$

Remark. A slightly different version of this was first proved by deLeeuw **(1)**. We shall not use the results of **(8)**, but instead shall give a self-contained proof of the theorem.

Proof. We show firstly that the ϕ is an "onto" mapping. For each $p \in P$ suppose an element $\Omega^p \in B(N_p, M_p)$ chosen. We must prove the existence of an element $A \in B(N, M)$ such that $[A^p] = [\Omega^p]$, $p \in P$. Let $T = \mathrm{Hom}_R(N, M)$, and let us set

$$T_p = \mathrm{Hom}_{R_p}(N_p, M_p) = R_p \mathrm{Hom}_R(N, M) = R_p T$$

for each prime ideal p.

For each $p \in P$, we may choose an element $\pi \in p$ such that $\pi \notin p^2$, and such that π does not lie in any other prime ideal in the set P. Set

$$a = \prod_{p \in P} \pi^{\gamma(p)};$$

then $a \in R$, and for each $p \in P$ we may write

$$a = \pi^{\gamma(p)} d_p, \qquad d_p \in R, \qquad d_p = \text{unit in } R_p.$$

Define the integral ideal $\mathfrak{b}$ by

$$(a) = \mathfrak{b} \cdot \prod_{p \in P} p^{\gamma(p)}.$$

Then $\mathfrak{b}$ is not a multiple of any of the prime ideals in P.

We now make use of equation (8) to deduce that for each $p \in P$, there exists an element $u^p \in T_p$ such that

$$a \cdot \Omega_p^p = g u^p - u^p g, \qquad g \in G.$$

On the other hand, T is a finitely-generated R-module, so there exist elements $t_1, \ldots, t_r \in T$ such that

$$T = Rt_1 + \ldots + Rt_r,$$

whence

$$T_p = R_p t_1 + \ldots + R_p t_r.$$

We may therefore write (for $p \in P$)

$$u^p = \sum_{i=1}^{r} \beta_i^p t_i, \qquad \beta_i^p \in R_p.$$

Let us now choose $\alpha_1, \ldots, \alpha_r \in R$ such that

$$\alpha_i \equiv \beta_i^p \pmod{\pi^{2\gamma(p)} R_p}, \qquad p \in P, \qquad \alpha_i \equiv 0 \pmod{\mathfrak{b}}.$$

Set

$$t = a^{-1} \sum_{i=1}^{r} \alpha_i t_i \in KT,$$

and define $\Lambda \in \mathrm{Hom}_R(G, KT)$ by

$$\Lambda_g = gt - tg, \qquad g \in G.$$

We shall show that this is the desired Λ, that is, $\Lambda \in B(N, M)$, and $[\Lambda^p] = [\Omega^p]$ for $p \in P$. For $p \in P$ we have

$$a(\Omega_g^p - \Lambda_g) = gv^p - v^p g, \qquad g \in G,$$

where

$$v^p = u^p - at = \sum_{i=1}^{r} (\beta_i^p - \alpha_i) t_i.$$

From the way in which the α_i were chosen, we may therefore write

$$v^p = \pi^{2\gamma(p)} d_p w^p,$$

where $w^p \in T_p$, and thus

$$\Omega_g^p - \Lambda_g = \pi^{\gamma(p)}(gw^p - w^p g), g \in G.$$

This proves that for each $p \in P$,

$$\Omega^p - \Lambda^p \in \pi^{\gamma(p)} B(N_p, M_p) \subset B'(N_p, M_p),$$

and shows incidentally that

$$\Lambda_g \in T_p, \qquad g \in G, \qquad p \in P. \tag{9}$$

On the other hand, we note that for each prime ideal $q \notin P$, the elements $a^{-1}\alpha_1, \ldots, a^{-1}\alpha_r$ all lie in R_q, and hence

$$\Lambda_g \in T_q, \qquad g \in G.$$

Coupled with (9), this implies that

$$\Lambda_g \in \bigcap_{q'} T_{q'}, \qquad g \in G,$$

where q' ranges over all prime ideals. The above intersection is precisely T, and so $\Lambda \in \mathrm{Hom}_R(G, T)$. That (7) holds follows at once from the definition of Λ; consequently, $\Lambda \in B(N, M)$. This completes the proof that ϕ is "onto."

In order to show that ϕ is an isomorphism, let $\Omega \in B(N, M)$ be such that $\Omega^p \in B'(N_p, M_p)$ for all $p \in P$; we must show that $\Omega \in B'(N, M)$. Since $\Omega^p \in B'(N_p, M_p)$, there exists for each $p \in P$ an element $u^p \in T_p$ such that

$$\Omega_g^p = gu^p - u^p g, \qquad g \in G.$$

By the preceding construction (with $a = 1$), there exists $\Lambda \in B'(N, M)$ (since now $t \in T$) such that

$$\Lambda_g^p \equiv \Omega_g^p \pmod{\pi^{\gamma(p)} T_p}, \qquad g \in G.$$

Therefore

$$\Lambda - \Omega \in \mathfrak{g}B(N, M) \subset B'(N, M),$$

which shows that $\Omega \in B'(N, M)$.

COROLLARY. *If N, N^*, M, M^* are G-modules such that $N \vee N^*$ and $M \vee M^*$, then $C(N, M) \cong C(N^*, M^*)$ as R-modules.*

More generally, let

$$L_1 \supset L_2 \supset \ldots \supset L_k \supset (0)$$

be a set of G-modules such that each is an R-direct summand of its predecessor. Define $N_i = L_i/L_{i+1}$ to be the factor G-module. Then as above, every element of L_1 is uniquely representable as an ordered k-tuple $(n_1, \ldots, n_k)$ $n_i \in N_i$, where

$$(n_1, \ldots, n_k) + (n_1', \ldots, n_k') = (n_1 + n_1', \ldots, n_k + n_k'),$$
$$\alpha(n_1, \ldots, n_k) = (\alpha n_1, \ldots, \alpha n_k)$$

for $n_i, n_i' \in N_i$, $\alpha \in R$. The action of G on L_1 is given by

$$g(n_1, \ldots, n_k) = (gn_1, gn_2 + \Lambda_g^{12} n_1, \ldots, gn_k + \Lambda_g^{12} n_1 + \ldots + \Lambda_g^{k-1,k} n_{k-1}),$$

where each $\Lambda_g^{ij} \in \mathrm{Hom}_R(N_i, N_j)$, and where the R-homomorphisms $\Lambda^{ij}: g \to \Lambda_g^{ij}$ satisfy conditions analogous to (7). Let $B(N_1, \ldots, N_k)$ denote the set of systems $\{\Lambda^{ij}\}$ satisfying these conditions. We denote the module L_1 by the symbol $(N_1, \ldots, N_k; \{\Lambda^{ij}\})$.

3. Isomorphisms of modules. Throughout this section, we fix an A-module L' with a composition series.

$$L = L_1' \supset L_2' \supset \ldots \supset L_k' \supset (0),$$

and let $N_i' = L_i'/L_{i+1}'$. We assume here that N_i' and N_j' are not isomorphic for $i \neq j$, and further that K is a splitting field for A. For any $L \in S(L')$, the A-module KL will have a composition series

$$KL = L_1'' \supset L_2'' \supset \ldots \supset L_k'' \supset (0)$$

in which $L_i''/L_{i+1}'' \cong N_i'$. Setting $L_i = L_i'' \cap L$, we see that L_i is a G-submodule of L for which $KL_i = L_i''$. Furthermore, L_{i+1} is a pure R-submodule of L_i, and therefore (by **5**) is an R-direct summand of L_i. Put $N_i = L_i/L_{i+1}$; then $KN_i \cong N_i'$, and

$$L = (N_1, \ldots, N_k; \{\Lambda^{ij}\})$$

for some $\{\Lambda^{ij}\} \in B(N_1, \ldots, N_k)$.

LEMMA 1. *Let $M_i, N_i \in S(N_i')$, $1 \leqslant i \leqslant k$, and suppose that*

$$(M_1, \ldots, M_k; \{\Lambda^{ij}\}) \cong (N_1, \ldots, N_k; \{\Omega^{ij}\}).$$

Then $M_i \cong N_i$, $1 \leqslant i \leqslant k$.

Proof. (A modified version of this is given in **(2)**.) It suffices to prove that if $(N, M; \Lambda) \cong (\bar{N}, \bar{M}; \bar{\Lambda})$, where $KN \cong K\bar{N}$ and $KM \cong K\bar{M}$, and where KN and KM have no common irreducible constituent, then $M \cong \bar{M}$ and $N \cong \bar{N}$. Once this is established, a simple induction argument completes the proof.

Suppose that θ: $(N, M; \Lambda) \cong (\bar{N}, \bar{M}; \bar{\Lambda})$ is given by

$$\theta(n, m) = \theta(n, 0) + \theta(0, m) = (\theta_1(n), \nu(n)) + (\mu(m), \theta_2(m)),$$

where

$$\theta_1 \in \mathrm{Hom}_R(N, \bar{N}),\ \nu \in \mathrm{Hom}_R(N, \bar{M}),\ \mu \in \mathrm{Hom}_R(M, \bar{N}),\ \theta_2 \in \mathrm{Hom}_R(M, \bar{M}).$$

From $\theta g(n, m) = g\theta(n, m)$ we obtain at once

$$\text{(10.1,10.2)} \qquad \theta_1 g + \mu\Lambda_g = g\theta_1, \qquad \mu g = g\mu,$$

$$\text{(10.3,10.4)} \qquad \bar{\Lambda}_g\theta_1 + g\nu = \nu g + \theta_2\Lambda_g, \qquad \theta_2 g = \bar{\Lambda}_g\mu + g\theta_2.$$

From (10.2) we have $\mu \in \mathrm{Hom}_G(M, \bar{N})$, and hence $\mu = 0$, since by hypothesis KM and $K\bar{N}$ have no common irreducible constituents. Equations (10.1) and (10.4) then imply that $\theta_1 \in \mathrm{Hom}_G(N, \bar{N})$ and $\theta_2 \in \mathrm{Hom}_G(M, \bar{M})$. Since θ is an isomorphism of $(N, M; \Lambda)$ onto $(\bar{N}, \bar{M}; \bar{\Lambda})$, we find readily that $\theta_1 : N \cong \bar{N}$ and $\theta_2 : M \cong \bar{M}$.

LEMMA 2. *Let* $(N_1, \ldots, N_k; \{\Lambda^{ij}\})$ *and* $(N_1, \ldots, N_k; \{\Omega^{ij}\})$ *be G-isomorphic modules in* $S(L')$, *where* $N_i \in S(N_i')$. *Then there exist units* $\beta_1, \ldots, \beta_k \in \mathfrak{u}$, *and homomorphisms* $t_{ij} \in \mathrm{Hom}_R(N_i, N_j)$, *such that the isomorphism between these G-modules is given by*

$$(n_1, \ldots, n_k) \to (\beta_1 n_1, \beta_2 n_2 + t_{12} n_1, \ldots, \beta_k n_k + t_{1k} n_1 + \ldots + t_{k-1,k} n_{k-1}).$$

Proof. From the proof of the preceding lemma, we find that the isomorphism must be given by

$$(n_1, \ldots, n_k) \to (\theta_1 n_1, \theta_2 n_2 + t_{12} n_1, \ldots, \theta_k n_k + t_{1k} n_1 + \ldots + t_{k-1,k} n_{k-1}),$$

with each θ_i: $N_i \cong N_i$ and each $t_{ij} \in \mathrm{Hom}_R(N_i, N_j)$. Since KN_i is an absolutely irreducible A-module, θ_i must be given by scalar multiplication by a unit of R. This completes the proof.

If U, V are R-modules, and $f_1, f_2 \in \mathrm{Hom}_R(U, V)$, we shall often abbreviate the congruence $f_1 \equiv f_2 \pmod{\mathfrak{g}^a \mathrm{Hom}_R(U, V)}$ as $f_1 \equiv f_2 \pmod{\mathfrak{g}^a}$. A similar notation will be used for R_g-modules.

LEMMA 3. *Let* $M_1, \ldots, M_k$ *be G-modules, not necessarily irreducible, and let*

$$L = (M_1, \ldots, M_k; \{\Lambda^{ij}\}), \qquad \bar{L} = (M_1, \ldots, M_k; \{\Omega^{ij}\})$$

be G-modules for which

$$\Lambda^{ij} \equiv \Omega^{ij} \pmod{\mathfrak{g}^a}, \qquad 1 \leqslant i < j \leqslant k,$$

where n is a fixed integer $\geq k-1$. Then there exists a G-isomorphism $\theta: L \cong \bar{L}$ such that $\theta \equiv I \pmod{\mathfrak{g}^{n-k+1}}$, where $I: L \cong \bar{L}$ is the R-isomorphism given by $(m_1, \ldots, m_k) \to (m_1, \ldots, m_k)$.

Proof. The result is trivial for $k = 1$; let $k > 1$, and assume the result holds at $k-1$. Let us set

$$\Delta = (M_2, \ldots, M_k; \Lambda^{23}, \ldots, \Lambda^{k-1,k}),$$

$$\bar{\Delta} = (M_2, \ldots, M_k; \Omega^{23}, \ldots, \Omega^{k-1,k}).$$

From the induction hypothesis we deduce the existence of a G-isomorphism $\theta_1: \Delta \cong \bar{\Delta}$ such that

$$\theta_1 \equiv I \pmod{\mathfrak{g}^{n-k+2}}.$$

The map $(m_1, \delta) \to (m_1, \theta_1\delta)$, where $m_1 \in M_1$, $\delta \in \Delta$, then gives a G-isomorphism

$$\theta_2 : (M_1, \Delta; \Lambda^{12}, \ldots, \Lambda^{1k}) \cong (M_1, \bar{\Delta}; \bar{\Lambda}^{12}, \ldots, \bar{\Lambda}^{1k})$$

for some $(\bar{\Lambda}^{12}, \ldots, \bar{\Lambda}^{1k}) \in B(M_1, \bar{\Delta})$, and we have

$$\theta_2 \equiv I \pmod{\mathfrak{g}^{n-k+2}}.$$

Now set

$$\bar{\Lambda} = (\bar{\Lambda}^{12}, \ldots, \bar{\Lambda}^{1k}), \qquad \Omega = (\Omega^{12}, \ldots, \Omega^{1k}).$$

Then we see that both $\bar{\Lambda}$ and Ω are elements of $B(M_1, \bar{\Delta})$, and that $\bar{\Lambda} \equiv \Omega \pmod{\mathfrak{g}^{n-k+2}}$. By considering this congruence for the powers of the prime ideals dividing $\mathfrak{g}$, the method of proof of Theorem 1 shows the existence of an element $W \in \mathrm{Hom}_R(M_1, \bar{\Delta})$ such that

$$(\bar{\Lambda} - \Omega)_g = gW - Wg, \qquad g \in G,$$

and where, furthermore, $W \equiv 0 \pmod{\mathfrak{g}^{n-k+1}}$. The map $(m_1, \bar{\delta}) \to (m_1, \bar{\delta} - Wm_1)$ then yields a G-isomorphism $\theta_3 : (M_1, \bar{\Delta}; \Omega) \cong (M_1, \bar{\Delta}; \Lambda)$, where

$$\theta_3 \equiv I \pmod{\mathfrak{g}^{n-k+1}}.$$

Therefore

$$\theta_3^{-1}\theta_2 : (M_1, \Delta; \Lambda^{12}, \ldots, \Lambda^{1k}) \to (M_1, \bar{\Delta}; \Omega^{12}, \ldots, \Omega^{1k})$$

is a G-isomorphism of L onto $\bar{L}$ such that

$$\theta_3^{-1}\theta_2 \equiv I \pmod{\mathfrak{g}^{n-k+1}}.$$

4. Integral classes and genera for modules with two distinct constituents. Throughout this section, we suppose that L' is an A-module with two distinct irreducible constituents N' and M'; we assume again that K is a splitting field for A. Let $S(L')$ be partitioned into r_G classes under G-isomorphism, and into r_g genera. We shall obtain formulas for r_G and r_g.

LEMMA 4. *Let* $N \in S(N')$, $M \in S(M')$. *Then* $(N, M; \Lambda) \cong (N, M; \overline{\Lambda})$ *if and only if there exists* $\beta \in \mathfrak{u}$ *such that* $[\overline{\Lambda}] = \beta[\Lambda]$.

Proof. From Lemma 2 we deduce the existence of units $\beta_1, \beta_2 \in \mathfrak{u}$, and of $t \in \mathrm{Hom}_R(N, M)$, such that the isomorphism $(N, M; \Lambda) \cong (N, M; \overline{\Lambda})$ is given by $(n, m) \to (\beta_1 n, \beta_2 m + tn)$. This implies

$$\overline{\Lambda}_g = \beta_1^{-1}\beta_2 \Lambda_g + g(\beta_1^{-1}t) - (\beta_1^{-1}t)g, \qquad g \in G.$$

Setting $\beta = \beta_1^{-1}\beta_2$, we have $[\overline{\Lambda}] = \beta[\Lambda]$. Conversely, starting from such a relation, we may reverse the steps to obtain an isomorphism of the modules.

LEMMA 5. *Let* $N \in S(N')$, $M \in S(M')$. *Then* $(N, M; \Lambda) \vee (N, M; \overline{\Lambda})$ *if and only if there exists an element* $\alpha \in R$ *such that* $(\alpha) + \mathfrak{g} = R$ *and* $[\overline{\Lambda}] = \alpha[\Lambda]$.

Proof. Let $(N, M; \Lambda) \vee (N, M; \overline{\Lambda})$. As in the preceding proof, we deduce that for each $p \in P$, there exists an element α_p which is a unit in R_p such that the classes $[\Lambda^p]$ and $[\overline{\Lambda}^p]$ in $C(N_p, M_p)$ are related by

$$[\overline{\Lambda}^p] = \alpha_p[\Lambda^p].$$

Choose $\alpha \in R$ such that $\alpha \equiv \alpha_p \pmod{p^{\tau(p)}}$ for each $p \in P$; then $(\alpha) + \mathfrak{g} = R$ Furthermore, $(\alpha - \alpha_p)B(N_p, M_p) \subset B'(N_p, M_p)$, so that

$$\alpha[\Lambda^p] = \alpha_p[\Lambda^p], \qquad p \in P.$$

Therefore $[\overline{\Lambda}^p] = [\alpha\Lambda^p]$ for all $p \in P$, and so by Theorem 1 we have $[\overline{\Lambda}] = [\alpha\Lambda] = \alpha[\Lambda]$.

Suppose now that $S(N')$ splits into ν genera; according to **(7)**, each genus splits into h classes under G-isomorphism. Let us choose representatives of the $h\nu$ classes, say $\{N_j^i : 1 \leqslant i \leqslant \nu, 1 \leqslant j \leqslant h\}$, so that all the modules with the same subscript lie in the same genus. Likewise choose representatives $\{M_j^i : 1 \leqslant i \leqslant \mu, 1 \leqslant j \leqslant h\}$ of the $h\mu$ classes into which $S(M')$ splits. Let $(N, M; \Gamma) \in S(L')$, and suppose $N \vee N_1^i$, $M \vee M_1^j$. Then for each $p \in P$, there exists an element

$$\Omega^p \in B((N_1^i)_p, (M_1^j)_p)$$

such that

$$(N_p, M_p; \Gamma^p) \cong ((N_1^i)_p, (M_1^j)_p; \Omega^p)$$

as A_p-modules. By Theorem 1, there exists $\Lambda \in B(N_1^i, M_1^j)$ such that $[\Lambda^p] = [\Omega^p]$ for all $p \in P$. Therefore

$$(N, M; \Gamma)^p \cong (N_1^i, M_1^j; \Lambda)_p, \qquad p \in P,$$

and so

$$(N, M; \Gamma) \vee (N_1^i, M_1^j; \Lambda).$$

Hence, every module in $S(L')$ is in the same genus as $(N_1^i, M_1^j; \Lambda)$ for some choice of i and j and some $\Lambda \in B(N_1^i, M_1^j)$. Further,

$$(N_1^i, M_1^j; \Lambda) \vee (N_1^{i'}, M_1^{j'}; \Lambda')$$

implies, by the method of proof of Lemma 1, that $i = i'$ and $j = j'$. Let us set

$$H_{ij} = \{(N_1^i, M_1^j; \Lambda) : \Lambda \in B(N_1^i, M_1^j)\}, 1 \leqslant i \leqslant \nu, 1 \leqslant j \leqslant \mu, \tag{11}$$

and suppose that H_{ij} splits into r_{ij} genera. Then we have at once

$$r_g = \sum_{i,j} r_{ij}. \tag{12}$$

On the other hand, any module in $S(L')$ is G-isomorphic to $(N_\rho^i, M_\sigma^j; \Lambda)$ for some i, j, ρ, σ and Λ. Further, by Lemma 1, two such modules cannot be isomorphic unless they have the same set of indices i, j, ρ, σ. Let us set

$$S(i, \rho; j, \sigma) = \{(N_\rho^i, M_\sigma^j; \Lambda) : \Lambda \in B(N_\rho^i, M_\sigma^j)\},$$

and suppose that $S(i, \rho; j, \sigma)$ splits into $s(i, \rho; j, \sigma)$ classes. Then

$$r_G = \sum_{i,j,\rho,\sigma} s(i, \rho; j, \sigma).$$

However, Lemma 4 states that $(N_\rho^i, M_\sigma^j; \Lambda) \cong (N_\rho^i, M_\sigma^j; \overline{\Lambda})$ if and only if there exists $\beta \in \mathfrak{u}$ such that $[\overline{\Lambda}] = \beta[\Lambda]$. Furthermore, the Corollary to Theorem 1 shows that $C(N_\rho^i, M_\sigma^j)$ is (as R-module) independent of ρ and σ. Therefore $s(i, \rho; j, \sigma) = s(i, 1; j, 1)$ for all ρ and σ, and we have

$$r_G = h^2 \sum_{i,j} s_{ij}, \tag{13}$$

where $s_{ij} = s(i, 1; j, 1)$ is the number of classes into which H_{ij} splits.

Before proceeding with the calculation of r_{ij} and s_{ij}, it will be convenient to introduce some notations. For a non-zero ideal $\mathfrak{a}$ in R, let $\phi(\mathfrak{a})$ denote the number of residue classes in $R/\mathfrak{a}$ which are relatively prime to $\mathfrak{a}$. If $\mathfrak{a} + \mathfrak{b} = R$, then $\phi(\mathfrak{a}\mathfrak{b}) = \phi(\mathfrak{a})\phi(\mathfrak{b})$. Next, let $u(\mathfrak{a})$ denote the number of distinct residue classes in $(\mathfrak{u} + \mathfrak{a})/\mathfrak{a}$; of course, $u(\mathfrak{a})$ is a divisor of $\phi(\mathfrak{a})$. However, $u(\mathfrak{a})$ is not a multiplicative function of $\mathfrak{a}$, as is seen from the example where K is the rational field.

LEMMA 6. *Let $N \in S(N')$, $M \in S(M')$, and $H = \{(N, M; \Lambda) : \Lambda \in B(N, M)\}$. Suppose H splits into r genera and s classes. Let $d(\mathfrak{a})$ be the number of elements in $C(N, M)$ with order ideal $\mathfrak{a}$. Then*

$$r = \sum_{\mathfrak{a}} d(\mathfrak{a})/\phi(\mathfrak{a}), \qquad s = \sum_{\mathfrak{a}} d(\mathfrak{a})/u(\mathfrak{a}),$$

both sums extending over all divisors of $\mathfrak{g}$.

(The order ideal of an element $c \in C(N, M)$ is $\{\alpha \in R : \alpha c = 0\}$.)

Proof. Let us use the symbol $(N, M; c)$ to denote the collection of mutually isomorphic modules $\{(N, M; \Lambda): \Lambda \in c\}$, where $c \in C(N, M)$. By Lemma 4, $(N, M; c)$ and $(N, M; c')$ cannot lie in the same genus unless c and c' have the same order ideal. Consider the set of $d(\mathfrak{a})$ elements of $C(N, M)$ with given order ideal $\mathfrak{a}$. For a fixed c in this set, all those c' of the form αc, where $\alpha \in R$ is such that $(\alpha) + \mathfrak{g} = R$, will yield modules in the same genus as those obtained from c. But as α ranges over all elements of R for which $(\alpha) + \mathfrak{g} = R$, αc gives exactly $\phi(\mathfrak{a})$ distinct elements of $C(N, M)$. Therefore

$$r = \sum_{\mathfrak{a}} d(\mathfrak{a})/\phi(\mathfrak{a}).$$

A similar argument gives the formula for s.

Let $d_{\mathfrak{p}}(\mathfrak{p}^a)$ denote the number of elements in $C(N_{\mathfrak{p}}, M_{\mathfrak{p}})$ having order ideal $\mathfrak{p}^a$. Then

$$d_{\mathfrak{p}}(\mathfrak{p}^a) = \tau(\mathfrak{p}^a) - \tau(\mathfrak{p}^{a-1}),$$

where $\tau(\mathfrak{p}^a)$ denotes the number of elements of $C(N_{\mathfrak{p}}, M_{\mathfrak{p}})$ which are annihilated by $\mathfrak{p}^a$. From Theorem 1,

$$d(\mathfrak{a}) = \prod_{\mathfrak{p} \in P} d_{\mathfrak{p}}(\mathfrak{p}^{a(\mathfrak{p})}), \text{ where } \mathfrak{a} = \prod_{\mathfrak{p} \in P} \mathfrak{p}^{a(\mathfrak{p})}.$$

We may therefore write

$$r = \prod_{\mathfrak{p} \in P} \left\{ \sum_{a=0}^{\tau(\mathfrak{p})} d_{\mathfrak{p}}(\mathfrak{p}^a)/\phi(\mathfrak{p}^a) \right\},$$

which confirms the result in (7) that the number of genera is the product over all $\mathfrak{p} \in P$ of the number of classes into which $S(L')$ splits under $A_{\mathfrak{p}}$-isomorphism. The corresponding multiplicative formula for s fails to hold, because $u(\mathfrak{a})$ is not multiplicative.

Applying Lemma 6 to our original problem, we may summarize our result as follows.

THEOREM 2. *Let $N^1, \ldots, N^r$ be representatives of the genera into which $S(N')$ splits, and $M^1, \ldots, M^s$ representatives of the genera of $S(M')$. For each divisor $\mathfrak{a}$ of $\mathfrak{g}$, let $d_{ij}(\mathfrak{a})$ denote the number of elements in $C(N^i, M^j)$ having order ideal $\mathfrak{a}$. Then $S(L')$ splits into $r_{\mathfrak{g}}$ genera and r_G classes, where*

$$r_{\mathfrak{g}} = \sum_{\mathfrak{a}} \sum_{i,j} d_{ij}(\mathfrak{a})/\phi(\mathfrak{a}), \quad r_G = h^2 \sum_{\mathfrak{a}} \sum_{i,j} d_{ij}(\mathfrak{a})/u(\mathfrak{a}).$$

Here, $\phi(\mathfrak{a})$ is the number of residue classes in $R/\mathfrak{a}$ which are relatively prime to $\mathfrak{a}$, and $u(\mathfrak{a})$ is the number of distinct elements of $(\mathfrak{u} + \mathfrak{a})/\mathfrak{a}$.

COROLLARY. *We have $r_G \geqslant h^2 r_{\mathfrak{g}}$, with equality provided that $\phi(\mathfrak{g}) = u(\mathfrak{g})$. Furthermore, if any $C(N^i, M^j)$ contains an element of order ideal $\mathfrak{a}$, where $u(\mathfrak{a}) < \phi(\mathfrak{a})$, then $r_G > h^2 r_{\mathfrak{g}}$.*

5. Integral classes and genera in the general case. Now let L' be an A-module with k distinct irreducible constituents, and let K be a splitting field for A. We preserve the notation introduced at the beginning of § 3. In this section we shall generalize the results given in the Corollary to Theorem 2.

For each κ $(1 \leqslant \kappa \leqslant k)$, let $\{N_\kappa^{ij}: 1 \leqslant i \leqslant \nu(\kappa), 1 \leqslant j \leqslant h\}$ be a full set of representatives of the $h\nu(\kappa)$ classes into which the set $\mathrm{S}(N_\kappa')$ splits; suppose these representative modules are so chosen that modules with the same indices i and κ lie in the same genus. Then every module in $S(L')$ is of the form

$$(N_1^{i_1 j_1}, \ldots, N_k^{i_k j_k}; \{\Lambda^{ij}\}).$$

Let $S(i_1, j_1; \ldots; i_k, j_k)$ be the set of all such modules obtained by letting $\{\Lambda^{ij}\}$ range over all systems in

$$B(N_1^{i_1 j_1}, \ldots, N_k^{i_k j_k}),$$

and let this set split into $r(i_1, j_1; \ldots; i_k, j_k)$ genera and $s(i_1, j_1; \ldots; i_k, j_k)$ classes. From the Corollary to Theorem 1, we see that $r(i_1, j_1; \ldots; i_k, j_k)$ is independent of $(j_1, \ldots, j_k)$, and therefore

$$r_g = h^{-k} \sum r(i_1, j_1, ; \ldots ; i_k, j_k), \quad r_G = \sum s(i_1, j_1, ; \ldots ; i_k, j_k),$$

both summations extending over all possible values of the i's and j's. This implies the result that

$$r_G \geqslant h^k r_g.$$

Finally, we prove:

THEOREM 3. *If* $u(\mathfrak{g}^{k-1}) = \phi(\mathfrak{g}^{k-1})$, *then* $r_G = h^k r_g$.

Proof. We remark that the hypothesis of the Theorem is simply a restatement of condition (3) given in the introduction. To prove the theorem, we need only show that $r(i_1, j_1; \ldots; i_k, j_k) = s(i_1, j_1; \ldots; i_k, j_k)$. We simplify the notation by letting $M_\kappa \in \mathrm{S}(N_\kappa')$, $1 \leqslant \kappa \leqslant k$. We shall prove that if

$$L = (M_1, \ldots, M_k; \{\Lambda^{ij}\}), \qquad \bar{L} = (M_1, \ldots, M_k; \{\bar{\Lambda}^{ij}\})$$

are such that $L \vee \bar{L}$, then also $L \cong \bar{L}$.

Since $L_p \cong \bar{L}_p$ for each $p \in P$, Lemma 2 shows the existence of units $\beta_1^p, \ldots, \beta_k^p$ in R_p, and homomorphisms

$$t_{ij}^p \in \mathrm{Hom}_p\,((M_i)_p, (M_j)_p)$$

such that the isomorphism $L_p \cong \bar{L}_p$ is given by

$$(m_1, \ldots, m_k) \to (\beta_1^p m_1, \beta_2^p m_2 + t_{12}^p m_1, \ldots, \beta_k^p m_k + t_{1k}^p m_1 + \ldots + t_{k-1,k}^p m_{k-1}).$$

By the hypothesis of the theorem, we may choose units $\beta_1, \ldots, \beta_k \in \mathfrak{u}$ such that

$$\beta_\kappa \equiv \beta_\kappa^p \pmod{p^{(k-1)\gamma(p)}}, \qquad p \in P, \qquad 1 \leqslant \kappa \leqslant k.$$

As in the proof of Theorem 1, we may choose homomorphisms $w_{ij} \in \mathrm{Hom}_R(M_i, M_j)$ such that

$$w_{ij}^p \equiv t_{ij}^p \mod p^{(k-1)\gamma(p)}, \qquad 1 \leqslant i < j \leqslant k, \qquad p \in P.$$

Then the map

$$(m_1, \ldots, m_k) \to (\beta_1 m_1, \beta_2 m_2 + w_{12} m_1, \ldots, \beta_k m_k + w_{1k} m_1 + \ldots + w_{k-1,k} m_{k-1})$$

gives a G-isomorphism of L onto a module L^* where $L^* = (M_1, \ldots, M_k; \{\Omega^{ij}\})$ and $\Omega^{ij} \equiv \overline{\Lambda}^{ij} \pmod{\mathfrak{g}^{k-1}}$ for $1 \leqslant i < j \leqslant k$. By Lemma 3 we then have $L^* \cong \overline{L}$, which completes the proof of the theorem.

It would be of interest to obtain formulas for r_G and $r_\mathfrak{g}$ which generalize those given in Theorem 2.

References

1. K. deLeeuw, *Some applications of cohomology to algebraic number theory and group representations*, unpublished.
2. F. E. Diederichsen, *Ueber die Ausreduktion ganzzahliger Gruppendarstellungen bei arithmetischer Äquivalenz*, Hamb. Abh., *14* (1938), 357–412.
3. D. G. Higman, *On orders in separable algebras*, Can. J. Math., *7* (1955), 509–515.
4. N. Jacobson, *The theory of rings* (New York, 1943).
5. I. Kaplansky, *Modules over Dedekind rings and valuation rings*, Trans. Amer. Math. Soc., *7* (1952), 327–40.
6. J.-M. Maranda, *On p-adic integral representations of finite groups*, Can. J. Math., *5* (1953), 344–355.
7. ——— *On the equivalence of representations of finite groups by groups of automorphisms of modules over Dedekind rings*, Can. J. Math., *7* (1955), 516–526.
8. I. Reiner, *Maschke modules over Dedekind rings*, Can. J. Math., *8* (1956), 329–334.
9. H. Zassenhaus, *Neuer Beweis der Endlichkeit der Klassenzahl bei unimodularer Äquivalenz endlicher ganzzahliger Substitutionsgruppen*, Hamb. Abh., *12* (1938), 276–288.

University of Illinois

WILEY BOOKS

INFORMATION THEORY AND STATISTICS

By Solomon Kullback, *The George Washington University*. 1959. 395 pages. $12.50.

A MODERN INTRODUCTION TO COLLEGE MATHEMATICS

By Israel H. Rose, *University of Massachusetts*. 1959. 530 pages. $6.50.

THE OFFICE SUPERVISOR: His Relations to Persons and to Work. Third Edition.

By Henry E. Niles, *The Baltimore Life Insurance Company*; Mary C. Niles, *Management Writer and Consultant*; and James C. Stephens, *U.S. Dept. of Labor*. 1959. 310 pages. $5.95.

PLANE TRIGONOMETRY

By A. W. Goodman, *University of Kentucky*. 1959. With tables: 267 pages, $4.50; Without tables: 197 pages, $3.75.

INTRODUCTORY COLLEGE MATHEMATICS

By Thomas L. Wade, *Florida State University*. 1959. 319 pages. $5.50.

OPERATIONS RESEARCH: Methods and Problems

By Maurice Sasieni, *Case Institute of Technology*; Arthur Yaspan, *Standard Oil Company (Ohio)*; and Lawrence Friedman, *M & M's Candies*. 1959. 316 pages. $8.50.

TESTING STATISTICAL HYPOTHESES

By E. L. Lehmann, *University of California, Berkeley*. 1959. 369 pages. $11.00.

METHODS OF CORRELATION AND REGRESSION ANALYSIS: Linear and Curvilinear. Third Edition

By M. Ezekiel, *Food and Agriculture Organization of the United Nations*; and Karl A. Fox, *Iowa State College*. 1959. 548 pages. Prob. $12.50.

MATHEMATICAL PROGRAMMING AND ELECTRICAL NETWORKS

By Jack B. Dennis, *Massachusetts Institute of Technology*. A Technology Press Research Monograph. 1959. 186 pages. $4.50.

REGRESSION ANALYSIS

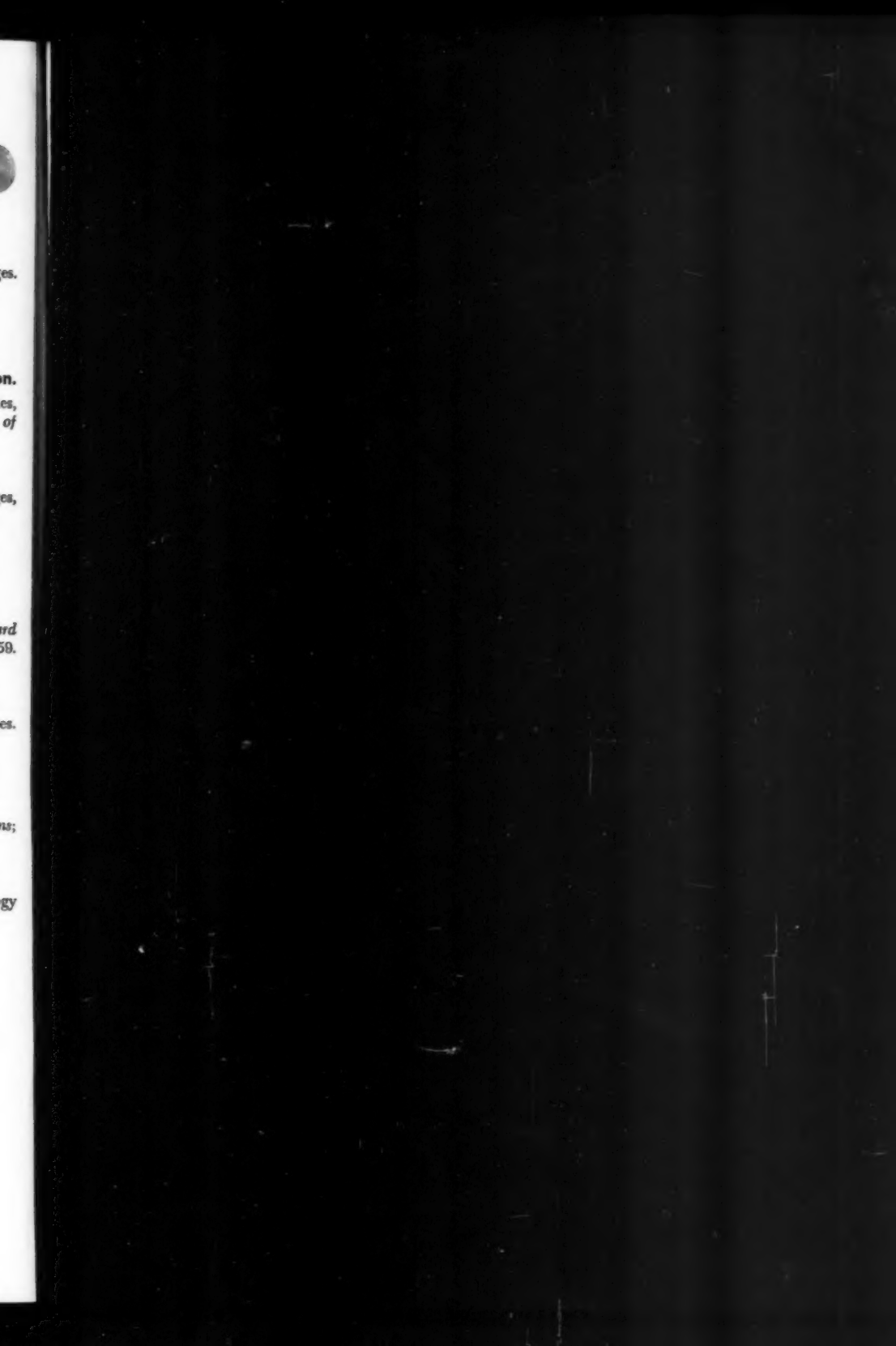
By E. J. Williams, *University of North Carolina*. 1959. In press.

COLLEGE ARITHMETIC

By W. I. Layton, *Stephen Austin State College*. 1959. 200 pages. $3.50.

SCIENTIFIC RUSSIAN

By George E. Condoyannis, *St. Peter's College*. 1959. 228 pages. $3.50.

Send for examination copies.

University of Toronto Press, Toronto, Ontario